

STIC-Biotech/ChemLib

168297

From: Chan, Christina
Sent: Tuesday, October 11, 2005 1:40 PM
To: Dunston, Jennifer; STIC-Biotech/ChemLib
Subject: RE: Sequence Search 10/826523

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

RECEIVED
OCT 11 2005
STIC/CHEN, Jennifer
(STIC)

-----Original Message-----

From: Dunston, Jennifer
Sent: Tuesday, October 11, 2005 1:19 PM
To: Chan, Christina
Subject: Sequence Search 10/826523

Please RUSH this search. It is for an after final amendment.

Please do a sequence search for the nucleic acid sequence of SEQ ID NO: 40 against the commercial and interference nucleotide databases.

The length of SEQ ID NO: 40 is 707 nucleotides.

Jennifer Dunston, Ph.D.
USPTO Art Unit 1636
REM 2B76
Mailbox: REM 2C70
(571) 272-2916

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 22:41:13 ; Search time 5848 Seconds
(without alignments)
847.469 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggaatccatgcgtcaatttt.....aaaatgacgcatggatcc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 8765912 seqs, 3504951483 residues

Total number of hits satisfying chosen parameters: 17531824

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	707	100.0	707	US-10-001-189-40
2	707	100.0	4613	Sequence 40, Appl Sequence 46, Appl
3	707	100.0	8999	Sequence 48, Appl Sequence 48, Appl
4	707	100.0	9012	Sequence 49, Appl Sequence 49, Appl
5	707	100.0	9013	Sequence 50, Appl Sequence 50, Appl
6	685	96.9	4943	Sequence 54, Appl Sequence 54, Appl
7	685	96.9	4944	Sequence 55, Appl Sequence 55, Appl

8	685	96.9	4944	13	US-10-001-189-56	Sequence 56, Appl
9	580.6	82.1	4164	22	US-10-220-335-158	Sequence 158, Appl
c 10	566.2	80.1	26565	16	US-10-085-959-91	Sequence 91, Appl
c 11	566.2	80.1	46819	14	US-10-114-170-72	Sequence 72, Appl
c 12	561.4	79.4	22306	14	US-10-114-170-251	Sequence 251, Appl
c 13	519.2	73.4	3662	13	US-10-001-189-41	Sequence 41, Appl
c 14	511	72.3	4941	13	US-10-001-189-53	Sequence 53, Appl
c 15	511	72.3	4951	13	US-10-001-189-51	Sequence 51, Appl
c 16	511	72.3	4952	13	US-10-001-189-52	Sequence 52, Appl
c 17	425.8	60.2	1427	20	US-10-363-345A-37517	Sequence 37517, A
c 18	425.8	60.2	1427	20	US-10-363-345A-37518	Sequence 37518, A
c 19	425.8	60.2	1427	21	US-10-363-483A-37517	Sequence 37517, A
c 20	425.8	60.2	1427	21	US-10-363-483A-37518	Sequence 37518, A
c 21	392.6	55.5	3622	22	US-10-450-763-30289	Sequence 30289, A
c 22	382.6	54.1	1427	20	US-10-363-345A-37519	Sequence 37519, A
c 23	382.6	54.1	1427	21	US-10-363-345A-37520	Sequence 37520, A
c 24	382.6	54.1	1427	21	US-10-363-483A-37519	Sequence 37519, A
c 25	382.6	54.1	1427	21	US-10-363-483A-37520	Sequence 37520, A
c 26	223	31.5	34063	14	US-10-114-170-96	Sequence 96, Appl
c 27	222.2	31.4	72480	17	US-10-418-837-2	Sequence 2, Appl
c 28	171.2	24.2	822	22	US-10-450-763-15978	Sequence 15978, A
c 29	96	13.6	847	10	US-09-798-889-27	Sequence 27, Appl
c 30	96	13.6	847	18	US-10-633-680-27	Sequence 27, Appl
c 31	78	11.0	9984	13	US-10-001-189-63	Sequence 63, Appl
c 32	72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, Appl
c 33	64.6	9.1	1275	22	US-10-450-763-23043	Sequence 23043, A
c 34	40.8	5.8	1523	22	US-10-450-763-15121	Sequence 15121, A
c 35	40	5.7	43	13	US-10-001-189-19	Sequence 19, Appl
c 36	40	5.7	3089	20	US-10-425-115-49726	Sequence 49726, A
c 37	39.8	5.6	676	20	US-10-425-115-91844	Sequence 91844, A
c 38	39.8	5.6	1673	18	US-10-425-114-21174	Sequence 21174, A
c 39	39.8	5.6	1796	18	US-10-425-114-23053	Sequence 23053, A
c 40	39.8	5.6	1872	20	US-10-425-115-91842	Sequence 91842, A
c 41	39	5.5	39	13	US-10-001-189-16	Sequence 16, Appl
c 42	38.4	5.4	1473	17	US-10-282-122A-20134	Sequence 20134, A
c 43	36.8	5.2	4187	19	US-10-437-963-91699	Sequence 91699, A
c 44	36.2	5.1	549	21	US-10-660-811A-23	Sequence 23, Appl
c 45	36.2	5.1	1403	19	US-10-767-701-13724	Sequence 13724, A

ALIGNMENTS

RESULT 1
US-10-001-189-40
; Sequence 40, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSCRIPTION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-32098
; CURRENT APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR FILING DATE: 2001-10-30
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge
; OTHER INFORMATION: sequence
US-10-001-189-40

Query Match 100.0%; Score 707; DB 13; Length 707;

```

Best Local Similarity 100.0%; Pred. No. 2.le-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 1 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60

Qy 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120

Qy 121 TCGGGAGGAAGAACCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 121 TCGGGAGGAAGAACCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180

Qy 181 CCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAGAACGCACTTTACCATGATGATTC 240
Db 181 CCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAGAACGCACTTTACCATGATGATTC 240

Qy 241 GGGAGGTTGTGCCATGACAGCCTTTTAAACGGTGAACCTGTTCCGAGCCACCTGGGATA 300
Db 241 GGGAGGTTGTGCCATGACAGCCTTTTAAACGGTGAACCTGTTCCGAGCCACCTGGGATA 300

Qy 301 CCAGTTTCGTCGGGCTTTTCCGGACACAGTTCCGGATGCTCAGCCGAGCGCATCAGCA 360
Db 301 CCAGTTTCGTCGGGCTTTTCCGGACACAGTTCCGGATGCTCAGCCGAGCGCATCAGCA 360

Qy 361 ACCGGAACAATACCGGCGACAGCGGAACTGCGGTGCGGTTGAGCATTAATGACAGCG 420
Db 361 ACCGGAACAATACCGGCGACAGCGGAACTGCGGTGCGGTTGAGCATTAATGACAGCG 420

Qy 421 GTGGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCGCAGAAAT 480
Db 421 GTGGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCGCAGAAAT 480

Qy 481 GGACATGATATCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 540
Db 481 GGACATGATATCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 540

Qy 541 CCGTGAGGACGGGAGACTGCGGTGCAATATGTTTTTACAGCGTGATGAGCGATGA 600
Db 541 CCGTGAGGACGGGAGACTGCGGTGCAATATGTTTTTACAGCGTGATGAGCGATGA 600

Qy 601 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATATCATATTG 660
Db 601 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATATCATATTG 660

Qy 661 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 707
Db 661 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 707

```

```

RESULT 2
US-10-001-189-46
; Sequence 46, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46

```

```

; LENGTH: 4613
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCR11-ITR
; OTHER INFORMATION: sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(922)
; US-10-001-189-46

Query Match 100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred. No. 5.6e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 294 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353

Qy 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 354 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 413

Qy 121 TCGGGAGGAAGAACCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 414 TCGGGAGGAAGAACCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 473

Qy 181 CCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAGAACGCACTTTACCATGATGATTC 240
Db 474 CCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAGAACGCACTTTACCATGATGATTC 533

Qy 241 GGGAGGTTGTGCCATGACAGCCTTTTAAACGGTGAACCTGTTCCGAGCCACCTGGGATA 300
Db 534 GGGAGGTTGTGCCATGACAGCCTTTTAAACGGTGAACCTGTTCCGAGCCACCTGGGATA 593

Qy 301 CCAGTTTCGTCGGGCTTTTCCGGACACAGTTCCGGATGCTCAGCCGAGCGCATCAGCA 360
Db 594 CCAGTTTCGTCGGGCTTTTCCGGACACAGTTCCGGATGCTCAGCCGAGCGCATCAGCA 653

Qy 361 ACCGGAACAATACCGGCGACAGCGGAACTGCGGTGCGGTTGAGCATTAATGACAGCG 420
Db 654 ACCGGAACAATACCGGCGACAGCGGAACTGCGGTGCGGTTGAGCATTAATGACAGCG 713

Qy 421 GTGGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCGCAGAAAT 480
Db 714 GTGGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCGCAGAAAT 773

Qy 481 GGACATGATATCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 540
Db 774 GGACATGATATCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 833

Qy 541 CCGTGAGGACGGGAGACTGCGGTGCAATATGTTTTTACAGCGTGATGAGCGATGA 600
Db 834 CCGTGAGGACGGGAGACTGCGGTGCAATATGTTTTTACAGCGTGATGAGCGATGA 893

Qy 601 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATATCATATTG 660
Db 894 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATATCATATTG 953

Qy 661 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 707
Db 954 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 1000

RESULT 3
US-10-001-189-48/c
; Sequence 48, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

```



```

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 8999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-48

Query Match 100.0%; Score 707; DB 13; Length 8999;
Best Local Similarity 100.0%; Pred. No. 7.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATATCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 8950 GATATCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8891

Qy 61 ATCATATCGTGGGTCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGCA 120
Db 8890 ATCATATCGTGGGTCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGCA 8831

Qy 121 TCGGGAGGAGAAAGCCGCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 8830 TCGGGAGGAGAAAGCCGCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 8771

Qy 181 CCGAGGATGACTGCTGCTGATGACGTTGAGCGGAAACCGACGTTTACCATGATGATTC 240
Db 8770 CCGAGGATGACTGCTGCTGATGACGTTGAGCGGAAACCGACGTTTACCATGATGATTC 8711

Qy 241 GGGAGGTTGGCCATGACGCTTTTAAACGTTGAACTGTTCTGTCAGGCCACCTGGGATA 300
Db 8710 GGGAGGTTGGCCATGACGCTTTTAAACGTTGAACTGTTCTGTCAGGCCACCTGGGATA 8651

Qy 301 CCAAGTTGCTCGGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCCGAAGCGCATCAGCA 360
Db 8650 CCAAGTTGCTCGGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCCGAAGCGCATCAGCA 8591

Qy 361 ACCCGAACAATACCGGCGACAGCGGAACTCCCGTGGCGGTGTCAGATTAAATGACAGCG 420
Db 8590 ACCCGAACAATACCGGCGACAGCGGAACTCCCGTGGCGGTGTCAGATTAAATGACAGCG 8531

Qy 421 GTGGGGCTGGGATATTAGTCAGCGAGGACGGGTATCTTGGTGGATCGCGCAGAAAT 480
Db 8530 GTGGGGCTGGGATATTAGTCAGCGAGGACGGGTATCTTGGTGGATCGCGCAGAAAT 8471

Qy 481 GGCATGATACCCCGTGGATTACCCGGCGGCGCTGTTTCATTACGTTTTCGAAAC 540
Db 8470 GGCATGATACCCCGTGGATTACCCGGCGGCGCTGTTTCATTACGTTTTCGAAAC 8411

Qy 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGCGTGTAGGAGCAGATGA 600
Db 8410 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGCGTGTAGGAGCAGATGA 8351

Qy 601 AGATGCTCGACACGCTCGAGAACACCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTG 660
Db 8350 AGATGCTCGACACGCTCGAGAACACCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTG 8291

Qy 661 TGAGTACGTTAAAGATAATCATCGGTAAATTTGACCATGGGATCC 707
Db 8290 TGAGTACGTTAAAGATAATCATCGGTAAATTTGACCATGGGATCC 8244

```

```

RESULT 4
US-10-001-189-49/c
; Sequence 49, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 9012
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
; OTHER INFORMATION: sequence
US-10-001-189-49

Query Match 100.0%; Score 707; DB 13; Length 9012;
Best Local Similarity 100.0%; Pred. No. 7.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATATCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 8963 GATATCCATCGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8904

Qy 61 ATCATATCGTGGGTCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGCA 120
Db 8903 ATCATATCGTGGGTCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGCA 8844

Qy 121 TCGGGAGGAGAAAGCCGCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 8843 TCGGGAGGAGAAAGCCGCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 8784

Qy 181 CCGAGGATGACTGCTGCTGATGACGTTGAGCGGAAACCGACGTTTACCATGATGATTC 240
Db 8783 CCGAGGATGACTGCTGCTGATGACGTTGAGCGGAAACCGACGTTTACCATGATGATTC 8724

Qy 241 GGGAGGTTGGCCATGACGCTTTTAAACGTTGAACTGTTCTGTCAGGCCACCTGGGATA 300
Db 8723 GGGAGGTTGGCCATGACGCTTTTAAACGTTGAACTGTTCTGTCAGGCCACCTGGGATA 8664

Qy 301 CCAAGTTGCTCGGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCCGAAGCGCATCAGCA 360
Db 8663 CCAAGTTGCTCGGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCCGAAGCGCATCAGCA 8604

Qy 361 ACCCGAACAATACCGGCGACAGCGGAACTCCCGTGGCGGTGTCAGATTAAATGACAGCG 420
Db 8603 ACCCGAACAATACCGGCGACAGCGGAACTCCCGTGGCGGTGTCAGATTAAATGACAGCG 8544

Qy 421 GTGGGGCTGGGATATTAGTCAGCGAGGACGGGTATCTTGGTGGATCGCGCAGAAAT 480
Db 8543 GTGGGGCTGGGATATTAGTCAGCGAGGACGGGTATCTTGGTGGATCGCGCAGAAAT 8484

Qy 481 GGCATGATACCCCGTGGATTACCCGGCGGCGCTGTTTCATTACGTTTTCGAAAC 540
Db 8483 GGCATGATACCCCGTGGATTACCCGGCGGCGCTGTTTCATTACGTTTTCGAAAC 8424

Qy 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGCGTGTAGGAGCAGATGA 600
Db 8423 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGCGTGTAGGAGCAGATGA 8364

```

Qy 601 AGATGCTCGACACGCTGCAGAACACGCGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 8363 AGATGCTCGACACGCTGCAGAACACGCGCTAGATTAAACCTAGAAAGATAATCATATTG 8304
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATAATTGACGCGATGGATCC 707
Db 8303 TGACGTACGTTAAAGATAATCATGCGTAAATAATTGACGCGATGGATCC 8257

RESULT 5
US-10-001-189-50/c
; Sequence 50, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 9013
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EGFP
; OTHER INFORMATION: sequence
US-10-001-189-50

Query Match 100.0%; Score 707; DB 13; Length 9013;
Best Local Similarity 100.0%; Pred. No. 7.9e-235; Mismatches 0; Indels 0; Gaps 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 8964 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8905

Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 8904 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8845

Qy 121 TCGGGGAGGAAGAGCCGCTGCTTTTCCCGGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
Db 8844 TCGGGGAGGAAGAGCCGCTGCTTTTCCCGGAGGTTGAAGCGGCATGGAAAGAGTTTG 8785

Qy 181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAACCGCACGCTTTACCATGATGATTC 240
Db 8784 CCGAGGATGACTGCTGCTGCATTGAGCTTGAGCGAAACCGCACGCTTTACCATGATGATTC 8725

Qy 241 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGACTGTTTTCAGGCCACCTGGGATA 300
Db 8724 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGACTGTTTTCAGGCCACCTGGGATA 8665

Qy 301 CCAGTTGTCGCGGCTTTTTCGGGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCA 360
Db 8664 CCAGTTGTCGCGGCTTTTTCGGGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCA 8605

Qy 361 ACCGGAACAATACCGGCGACAGCCGGAACCTGCGGTGCGGCTGTGACAGATTAAATGACAGCG 420
Db 8604 ACCGGAACAATACCGGCGACAGCCGGAACCTGCGGTGCGGCTGTGACAGATTAAATGACAGCG 8545

Qy 421 GTGCGGCGCTGGGATATTTAAGTACGCGAGGACGGGTATCTGCGCTGGATGCCGAGAAAT 480
Db 8544 GTGCGGCGCTGGGATATTTAAGTACGCGAGGACGGGTATCTGCGCTGGATGCCGAGAAAT 8485

Qy 481 GGACATGATACCCCGTAGTATTACCCGCGGGCGGCTCGTTTCATTCACTTTTGAAC 540
Db 8484 GGACATGATACCCCGTAGTATTACCCGCGGGCGGCTCGTTTCATTCACTTTTGAAC 8425

Qy 541 CCGTGGAGGACGCGGACAGCTCGCGGTGCAAAATGTTTTTACAGCTGATGGAGCAGATGA 600
Db 8424 CCGTGGAGGACGCGGACAGCTCGCGGTGCAAAATGTTTTTACAGCTGATGGAGCAGATGA 8365

Qy 601 AGATGCTCGACACGCTGCAGAACACCGCAATAGATTAAACCTAGAAAGATAATCATATTG 660
Db 8364 AGATGCTCGACACGCTGCAGAACACCGCAATAGATTAAACCTAGAAAGATAATCATATTG 8305

Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATAATTGACGCGATGGATCC 707
Db 8304 TGACGTACGTTAAAGATAATCATGCGTAAATAATTGACGCGATGGATCC 8258

RESULT 6
US-10-001-189-54
; Sequence 54, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 4943
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-ECFP
; OTHER INFORMATION: sequence
US-10-001-189-54

Query Match 96.9%; Score 685; DB 13; Length 4943;
Best Local Similarity 99.2%; Pred. No. 2.7e-227; Mismatches 0; Indels 6; Gaps 1;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy 121 TCGGGGAGGAAGAGCCGCTGCTTTTCCCGGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
Db 839 TCGGGGAGGAAGAGCCGCTGCTTTTCCCGGAGGTTGAAGCGGCATGGAAAGAGTTTG 898

Qy 181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAACCGCACGCTTTACCATGATGATTC 240
Db 899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAACCGCACGCTTTACCATGATGATTC 958

Qy 241 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGACTGTTTTCAGGCCACCTGGGATA 300
Db 959 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGACTGTTTTCAGGCCACCTGGGATA 1018

Qy 301 CCAGTTGTCGCGGCTTTTTCGGGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCA 360
Db 301 CCAGTTGTCGCGGCTTTTTCGGGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCA 360

Db 1019 CCAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGGTCAAGCCGAGCATCAGCA 1078
Qy 361 ACCCGAACATATACCGGCGACAGCGGAACTGCGGTGCGCGGTGTCAGATTAAATGACAGCG 420
Db 1079 ACCCGAACATATACCGGCGACAGCGGAACTGCGGTGCGCGGTGTCAGATTAAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTAGTCAGCGAGGACGCGGTATCTGCTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTAGTCAGCGAGGACGCGGTATCTGCTGGATCCGCGAGAAAT 1198
Qy 481 GGACATCGATACCCCGTGAGTTACCCGCGCGCGCGCTCGTTCATTTCAGCTTTTGAAC 540
Db 1199 GGACATCGATACCCCGTGAGTTACCCGCGCGCGCTCGTTCATTTCAGCTTTTGAAC 1252
Qy 541 CCGTGGAGGACGGGACAGCTCGGGTGCAGAACTGTTTTACAGCGTGATGGAGCAGATGA 600
Db 1253 CCGTGGAGGACGGGACAGCTCGGGTGCAGAACTGTTTTACAGCGTGATGGAGCAGATGA 1312
Qy 601 AGATGCTGACAGCTGCAGAACACGCGAGCTAGATTAACTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTGACAGCTGCAGAACACGCGAGCTAGATTAACTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 1419

RESULT 7

US-10-001-189-55
; Sequence 55, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
; OTHER INFORMATION: sequence
US-10-001-189-55

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.7e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 GGATCCCATCGTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCGTGGGCTCTTTTCCGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCA 120
Db 779 ATCATATCGTGGGCTCTTTTCCGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCA 838
Qy 121 TCGGGAGGAGGAGGACCGGCGCTTTTCCCGCAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 839 TCGGGAGGAGGAGGACCGGCGCTTTTCCCGCAGGTTGAAGCGGCATGGAAGAGTTTG 898
Qy 181 CCGAGGATGACTGCTGTCATTTGACGTTGAGCGGAAACGACAGCTTTTACCATGATGATC 240

Db 899 CCAGGATGACTGCTGTCATTTGAGCGTTGAGCGAAACGACGTTTACCATGATGATTC 958
Qy 241 GGAAGGTGTGGCCATGCGCTTTAAACGGTGAACGTTTCTGTTTACGGCCACCTGGGATA 300
Db 959 GGAAGGTGTGGCCATGCGCTTTAAACGGTGAACGTTTCTGTTTACGGCCACCTGGGATA 1018
Qy 301 CCAAGTTCTGTCGGCGCTTTTCCGGACACAGTTTCCGGATGGTCAAGCCGAGCGCATCAGCA 360
Db 1019 CCAAGTTCTGTCGGCGCTTTTCCGGACACAGTTTCCGGATGGTCAAGCCGAGCGCATCAGCA 1078
Qy 361 ACCGGAACATATACCGGCGACAGCGGAACTGCGGTGCGCGGTGTCAGATTAAATGACAGCG 420
Db 1079 ACCGGAACATATACCGGCGACAGCGGAACTGCGGTGCGCGGTGTCAGATTAAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTAGTCAGCGAGGACGCGGTATCTGCTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTAGTCAGCGAGGACGCGGTATCTGCTGGATCCGCGAGAAAT 1198
Qy 481 GGACATGATACCCCGTGAGTTACCCGCGCGCGCGCTCGTTCATTTCAGCTTTTGAAC 540
Db 1199 GGACATGATACCCCGTGAGTTACCCGCGCGCGCGCTCGTTCATTTCAGCTTTTGAAC 1252
Qy 541 CCGTGGAGGACGGGACAGCTCGGGTGCAGAACTGTTTTACAGCGTGATGGAGCAGATGA 600
Db 1253 CCGTGGAGGACGGGACAGCTCGGGTGCAGAACTGTTTTACAGCGTGATGGAGCAGATGA 1312
Qy 601 AGATGCTGACAGCTGCAGAACACGCGAGCTAGATTAACTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTGACAGCTGCAGAACACGCGAGCTAGATTAACTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 1419

RESULT 8

US-10-001-189-56
; Sequence 56, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-56

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.7e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 GGATCCCATCGTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778


```

Query Match      80.1%; Score 566.2; DB 16; Length 26565;
Best Local Similarity 97.8%; Pred. No. 1.9e-185;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db 3221 ATCCAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3162

Qy 105 TGGCGCTATCTGGGATCGGGAGGAAGAACCGTCCTTTTTCGGGAGGTGGAAGCG 164
Db 3161 TGGCGCTATCTGGGATCGGGAGGAAGAACCGTCCTTTTTCGGGAGGTGGAAGCG 3102

Qy 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAGC 224
Db 3101 GCATGGAAGAGTTTCCGAGGACGACTGTTGCTGCTATGACGTTGAGCGAAACGCAAGC 3042

Qy 225 TTACCATGATGATTCGGGAAGGTGTCGCCATGACGACGCTTTTAAACGTTGACTGTCGTT 284
Db 3041 TTACCATGATGATTCGGGAAGGTGTCGCCATGACGACGCTTTTAAACGTTGACTGTCGTT 2982

Qy 285 CAGCCACCTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTTCCGATGCTCAGC 344
Db 2981 CAGCCACCTGGGATACAGCTGTCGGCTTTTTCGGGACACAGTTTCCGATGCTCAGC 2922

Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTCCGTCGGGTGTG 404
Db 2921 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTCCGTCGGGTGTG 2862

Qy 405 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGGC 464
Db 2861 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGGC 2802

Qy 465 TGGATCCGCGAAGTACGATGATACCGCTGAGTTACCGCGCGCGCGCTCGCTTC 524
Db 2801 TGGATCCGCGAAGTACGATGATACCGCTGAGTTACCGCGCGCGCGCTCGCTTC 2742

Qy 525 ATTACAGCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 584
Db 2741 ATTACAGCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 2682

Qy 585 GTGATGAGCGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 631
Db 2681 GTGATGAGCGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 2635

```

RESULT 11

```

US-10-114-170-72
; Sequence 72, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Ferna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002

```

```

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46819
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-114-170-72

Query Match      80.1%; Score 566.2; DB 14; Length 46819;
Best Local Similarity 97.8%; Pred. No. 2.6e-185;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db 14795 ATCCAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 14854

Qy 105 TGGCGCTATCTGGGATCGGGAGGAAGAACCGTCCTTTTTCGGGAGGTGGAAGCG 164
Db 14855 TGGCGCTATCTGGGATCGGGAGGAAGAACCGTCCTTTTTCGGGAGGTGGAAGCG 14914

Qy 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAGC 224
Db 14915 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAGC 14974

Qy 225 TTACCATGATGATTCGGGAAGGTGTCGCCATGACGACGCTTTTAAACGTTGACTGTCGTT 284
Db 14975 TTACCATGATGATTCGGGAAGGTGTCGCCATGACGACGCTTTTAAACGTTGACTGTCGTT 15034

Qy 285 CAGGCCACCTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTTCCGATGCTCAGC 344
Db 15035 CAGGCCACCTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTTCCGATGCTCAGC 15094

Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTCCGTCGGGTGTG 404
Db 15095 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTCCGTCGGGTGTG 15154

Qy 405 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGGC 464
Db 15155 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGGC 15214

Qy 465 TGGATGCGCAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 524
Db 15215 TGGATGCGCAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 15274

Qy 525 ATTACAGCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 584
Db 15275 ATTACAGCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 15334

Qy 585 GTGATGAGCGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 631
Db 15335 GTGATGAGCGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 15381

RESULT 12
US-10-114-170-251/c
; Sequence 251, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:

```


Qy 362 CCCGAAACAATACCGCGGACAGCCGGAACCTGCGGTGTCAGATTAATGACAGCGG 421
 Db 1343 CCCGAAACAATACCGCGGACAGCCGGAACCTGCGGTGTCAGATTAATGACAGCGG 1402
 Qy 422 TCGCGGCTGGATATTACGTCAGCGAGGAGGATCTCTGGCTGGATGCGGAGAAATG 481
 Db 1403 TCGCGGCTGGATATTACGTCAGCGAGGAGGATCTCTGGCTGGATGCGGAGAAATG 1462
 Qy 482 GACATGATACCCCGTGGATTTACCGCGGCGCGCTCGTTCAATTCACGTT 533
 Db 1463 GACATGATACCCCGTGGATTTACCGCGGCGCGCTCGTTGGGTAAATCATGTT 1514

RESULT 14
 US-10-001-189-53/c
 ; Sequence 53, Application US/10001189
 ; Publication No. US20020173634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; TITLE OF INVENTION: VECTOR PIGGYBAC
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/001.189
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 53
 ; LENGTH: 4941
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECFP
 ; OTHER INFORMATION: sequence
 US-10-001-189-53

Query Match 72.3%; Score 511; DB 13; Length 4941;
 Best Local Similarity 100.0%; Pred. No. 1.5e-166;
 Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 CATCGCTCAATTTTACGCGAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 66
 Db 1129 CATCGCTCAATTTTACGCGAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 1070
 Qy 67 TCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGATCGGG 126
 Db 1069 TCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCNAGCTGGCGTATCTGGGATCGGG 1010
 Qy 127 AGGAAGAGCCGCTGCTTTTTCGGGAGGTTGAAGCGGATGAAAGAGTTTCCGAGG 186
 Db 1009 AGGAAGAGCCGCTGCTTTTTCGGGAGGTTGAAGCGGATGAAAGAGTTTCCGAGG 950
 Qy 187 ATGACTGCTGCTGATTTGAGCGGAAACGCGAGTTTACCATGATGATTCGGGAAG 246
 Db 949 ATGACTGCTGCTGATTTGAGCGGAAACGCGAGTTTACCATGATGATTCGGGAAG 890
 Qy 247 GTGTGGCCATGACGCTTTTAAACGGTGAATCTGTTTTCAGGCCACCTGGGATACCAAGTT 306
 Db 889 GTGTGGCCATGACGCTTTTAAACGGTGAATCTGTTTTCAGGCCACCTGGGATACCAAGTT 830
 Qy 307 CGTGGCGGCTTTTTCGGGATGATTCGGGATGATTCGGGATGATTCGGGATACCAAGTT 366
 Db 829 GTGTGGCCATGACGCTTTTAAACGGTGAATCTGTTTTCAGGCCACCTGGGATACCAAGTT 770
 Qy 367 CGTGGCGGCTTTTTCGGGATGATTCGGGATGATTCGGGATGATTCGGGATACCAAGTT 426
 Db 829 CGTGGCGGCTTTTTCGGGATGATTCGGGATGATTCGGGATGATTCGGGATACCAAGTT 770
 Qy 367 ACAATACCGCGGACAGCCGGAACCTGCGGTGTCAGATTAATGACAGCGGTGCGG 426

Db 769 ACAATACCGCGGACAGCCGGAACCTGCGGTGTCAGATTAATGACAGCGGTGCGG 710
 Qy 427 CGCTGGGATATTACGTCAGCGAGGAGGATCTCTGGCTGGATGCGGCAAAATGGACAT 486
 Db 709 CGCTGGGATATTACGTCAGCGAGGAGGATCTCTGGCTGGATGCGGCAAAATGGACAT 650
 Qy 487 GATATACCCCGTGGATTTACCGCGGCGCGCGC 517
 Db 649 GATATACCCCGTGGATTTACCGCGGCGCGCGC 619

RESULT 15
 US-10-001-189-51/c
 ; Sequence 51, Application US/10001189
 ; Publication No. US20020173634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; TITLE OF INVENTION: VECTOR PIGGYBAC
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/001.189
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 51
 ; LENGTH: 4951
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-BYFP
 ; OTHER INFORMATION: sequence
 US-10-001-189-51

Query Match 72.3%; Score 511; DB 13; Length 4951;
 Best Local Similarity 100.0%; Pred. No. 1.5e-166;
 Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 CATCGCTCAATTTTACGCGAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 66
 Db 1129 CATCGCTCAATTTTACGCGAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 1070
 Qy 67 TCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGATCGGG 126
 Db 1069 TCGTGGGTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGGCGTATCTGGGATCGGG 1010
 Qy 127 AGGAAGAGCCGCTGCTTTTTCGGGAGGTTGAAGCGGATGAAAGAGTTTCCGAGG 186
 Db 1009 AGGAAGAGCCGCTGCTTTTTCGGGAGGTTGAAGCGGATGAAAGAGTTTCCGAGG 950
 Qy 187 ATGACTGCTGCTGATTTGAGCGGAAACGCGAGTTTACCATGATGATTCGGGAAG 246
 Db 949 ATGACTGCTGCTGATTTGAGCGGAAACGCGAGTTTACCATGATGATTCGGGAAG 890
 Qy 247 GTGTGGCCATGACGCTTTTAAACGGTGAATCTGTTTTCAGGCCACCTGGGATACCAAGTT 306
 Db 889 GTGTGGCCATGACGCTTTTAAACGGTGAATCTGTTTTCAGGCCACCTGGGATACCAAGTT 830
 Qy 307 CGTGGCGGCTTTTTCGGGATGATTCGGGATGATTCGGGATGATTCGGGATACCAAGTT 366
 Db 829 GTGTGGCCATGACGCTTTTAAACGGTGAATCTGTTTTCAGGCCACCTGGGATACCAAGTT 770
 Qy 367 ACAATACCGCGGACAGCCGGAACCTGCGGTGTCAGATTAATGACAGCGGTGCGG 426
 Db 769 ACAATACCGCGGACAGCCGGAACCTGCGGTGTCAGATTAATGACAGCGGTGCGG 710
 Qy 427 CGCTGGGATATTACGTCAGCGAGGAGGATCTCTGGCTGGATGCGGCAAAATGGACAT 486

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 13:49:25 ; Search time 26459 Seconds
(without alignments)
1294.752 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatccatgcgtcaatttt.....aaattgacgcgtggatcc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.4	82.8	20067	12 XXU02425	U02425 Cloning vec
2	585.4	82.8	20125	12 XXU02426	U02426 Cloning vec
3	585.4	82.8	42529	12 CVU39284	U39284 Cloning vec
4	585.4	82.8	42530	12 CVU39285	U39285 Cloning vec
5	585.4	82.8	42531	12 CVU39286	U39286 Cloning vec
6	585.4	82.8	42704	12 CVU37692	U37692 Cloning vec
7	585.4	82.8	48502	7 LAMCG	J02459 Bacterioph
8	574.6	81.3	5969	12 AF434923	AF434923 Expressio
9	574.6	81.3	6182	12 AF434925	AF434925 Expressio
10	574.6	81.3	6182	12 AF434926	AF434926 Expressio
11	574.6	81.3	6182	12 AF434927	AF434927 Expressio
12	574.6	81.3	6182	12 AF434931	AF434931 Expressio
13	574.6	81.3	6182	12 AF434932	AF434932 Expressio
14	574.6	81.3	7421	12 AF434924	AF434924 Expressio
15	574.6	81.3	7634	12 AF434928	AF434928 Expressio
16	574.6	81.3	7634	12 AF434929	AF434929 Expressio
17	574.6	81.3	7634	12 AF434930	AF434930 Expressio
18	574.6	81.3	7634	12 AF434933	AF434933 Expressio
19	574.6	81.3	7634	12 AF434934	AF434934 Expressio

20	566.2	80.1	14164	1 AE005330	AE005330 Escherich
21	566.2	80.1	26565	6 AX702505	AX702505 Sequence
22	566.2	80.1	46819	6 AR204176	AR204176 Sequence
23	566.2	80.1	305325	1 AE016765	AE016765 Escherich
24	564.6	79.9	46897	6 BD184772	BD184772 Nucleic a
25	564.6	79.9	222605	1 AP002555	AP002555 Escherich
26	561.4	79.4	22306	6 AR204355	AR204355 Sequence
27	561.4	79.4	91772	6 BD184765	BD184765 Nucleic a
28	561.4	79.4	286485	1 AP002557	AP002557 Escherich
29	407	57.6	46375	7 AF064539	AF064539 Bacteriop
30	395.8	56.0	64137	2 AC101160	AC101160 Mus muscu
31	395.8	56.0	179840	2 AC068458	AC068458 Homo sapi
32	395.8	56.0	179871	2 AC024602	AC024602 Homo sapi
33	395.8	56.0	194028	2 AC111077	AC111077 Homo sapi
34	395.8	56.0	206486	2 AC074218	AC074218 Mus muscu
35	362.2	51.2	228516	2 AC008594	AC008594 Homo sapi
36	343.2	48.5	68028	2 AC101473	AC101473 Mus muscu
37	329.8	46.6	66450	2 AC060821	AC060821 Homo sapi
38	320.4	45.3	63579	2 AC100018	AC100018 Mus muscu
39	320.2	45.3	63690	2 AC125419	AC125419 Homo sapi
40	231.8	32.8	29466	1 AE008818	AE008818 Salmonell
41	223	31.5	8134	7 PH2HEADTL	M81255 Bacterioph
42	223	31.5	13264	1 AE005368	AE005368 Escherich
43	223	31.5	13386	1 AE015100	AE015100 Shigella
44	223	31.5	34063	6 AR204200	AR204200 Sequence
45	223	31.5	44029	6 BD184774	BD184774 Nucleic a

ALIGNMENTS

RESULT 1
XXU02425
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
ORIGIN
Query Match
Best Local Similarity
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XXU02425 20067 bp DNA linear SYN 29-JAN-1997
Cloning vector lambda EMBL3, left arm.
U02425
U02425.1 GI:413791
Cloning vector lambda EMBL3
Cloning vector lambda EMBL3
other sequences; artificial sequences; vectors.
1 (bases 1 to 20667)
Frischauf, A.M., Lehrach, H., Poustka, A. and Murray, N.
Lambda replacement vectors carrying polylinker sequences
J. Mol. Biol. 170 (4), 827-842 (1983)
84064856
6315951
2 (bases 1 to 20667)
Kitts, P.A.
CLONTECH Vectors On Disc version 1.3
Unpublished
3 (bases 1 to 20667)
Kitts, P.A.
Direct Submission
Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
This sequence has been compiled from information in the sequence
databases, published literature and other sources. If you suspect
there is an error in this sequence, please contact CLONTECH's
Technical Service Department at (415) 424-8222 or (800) 662-2566,
extension 3 or E-mail TECH@CLONTECH.COM.
1. .20067
Location/Qualifiers
/organism="Cloning vector lambda EMBL3"
/mol_type="genomic DNA"
/db_xref="taxon:11783"

QY

45 ATCTAGTCGTCATCAGGATCATATCGTCGGGTCTTTTTCGCGGTCTAGTCATCGCCCAAGC 104

[illegible]

RESULT 2	XXU02426	20125 bp	DNA	linear	SYN 29-JAN-1997
LOCUS	XXU02426				
DEFINITION	Cloning vector lambda EMBL3 SP6/T7, left arm.				
ACCESSION	U02426				
VERSION	U02426.1				
KEYWORDS	GI:413792				
SOURCE	Cloning vector lambda EMBL3 SP6/T7				
ORGANISM	Cloning vector lambda EMBL3 SP6/T7				
	other sequences; artificial sequences; vectors.				

AUTHORS	KITTS,P.A.
TITLE	CLONTECH Vectors On Disc version 1.3
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 20135)
AUTHORS	KITTS,P.A.
TITLE	Direct Submission
JOURNAL	Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 120 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT	This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.

FEATURES	source
Location/Qualifiers	
1. .20125	
/organism="Cloning vector lambda EMBL3 SP6/T7"	
/mol_type="genomic DNA"	
/db_xref="taxon:31784"	

Query Match	82.8%	Score 585.4;	DB 12;	Length 20125;
Best Local Similarity	99.8%	Pred. No. 2.3e-148;		

Matches	586;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCGGGTCACTCATCGCCCAAGC	104						
Db	3055	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCGGGTCACTCATCGCCCAAGC	3114						
Qy	105	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG	164						
Db	3115	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG	3174						
Qy	165	GCATGGAAGACAGTTTGCCTCAGGAGTACCTGCTGCTGCATTCAGCTTGAGCGAAACGCACG	224						
Db	3175	GCATGGAAGAGTTTGCCTCAGGAGTACCTGCTGCTGCATTCAGCTTGAGCGAAACGCACG	3234						
Qy	225	TTTATCCATGATGATTCCGGGAAGGTGTGGCCATGCACGCCCTTTAAACGGTGAATGTTCGTT	284						
Db	3235	TTTATCCATGATGATTCCGGGAAGGTGTGGCCATGCACGCCCTTTAAACGGTGAATGTTCGTT	3294						
Qy	285	CAGGCCACCTGGGATACCACTTCGTTCGGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC	344						
Db	3295	CAGGCCACCTGGGATACCACTTCGTTCGGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC	3354						
Qy	345	CCGAAGCGCATCAGCAACCCGAAACAATACCCGGCGACAGCCGCGAACTGCCGTGCCGGTGTG	404						
Db	3355	CCGAAGCGCATCAGCAACCCGAAACAATACCCGGCGACAGCCGCGAACTGCCGTGCCGGTGTG	3414						
Qy	405	CAGATTAAATCAGACCGGTGCGGCGCTGGGATATTAAGTACGAGGAGACGGGTATCTTGGC	464						
Db	3415	CAGATTAAATCAGACCGGTGCGGCGCTGGGATATTAAGTACGAGGAGACGGGTATCTTGGC	3474						
Qy	465	TGGATCCGCGAGAAATGGACAATGGATACCCCGTGAGTTTACCCGCGGGCGCGCTCGTTC	524						
Db	3475	TGGATCCGCGAGAAATGGACAATGGATACCCCGTGAGTTTACCCGCGGGCGCGCTCGTTC	3534						
Qy	525	ATTACACGTTTTTGAAACCCGTGGAAGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC	584						
Db	3535	ATTACACGTTTTTGAAACCCGTGGAAGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC	3594						
Qy	585	GTGATGAGCAGATGAAGATGCTCGACACCGTGCAGAAACACGCAGCT	631						
Db	3595	GTGATGAGCAGATGAAGATGCTCGACACCGTGCAGAAACACGCAGCT	3641						

RESULT	3
LOCUS	CVU39284
DEFINITION	CVU39284 42529 bp DNA linear SYN 13-APR-1986 Cloning vector TLF97-1, lambda phage lacZ translational fusion vector, complete sequence.
ACCESSION	U39284
VERSION	U39284.1
KEYWORDS	GI:1066304
SOURCE	. Cloning vector TLF97-1
ORGANISM	Cloning vector TLF97-1 other sequences: artificial sequences; vectors.

```

AUTHORS      St Pierre, R. and Linn, T.
TITLE        A refined vector system for the in vitro construction of
              single-copy transcriptional or translational fusions to lacZ
JOURNAL      Gene 169 (1), 65-68 (1996)
MEDLINE      96186904
PubMed      8635751
REFERENCE    2 (bases 1 to 42529)
AUTHORS      StPierre, R.
TITLE        Direct Submission
JOURNAL      Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
              University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
    source    1..42529
               /organism="Cloning vector TLF97-1"
               /mol_type="genomic DNA"
               /db_xref="taxon:43840"
               complement(19404..19489)
               /note="rrnB2 terminator"
terminator

```

[illegible]

gene	complement (19507. .20760) /gene="lacy" complement (19507. .20760) /gene="lacy" /codon_start=1 /transl_table=11 /product="lactose permease" /protein_id="AAC53651.1" /db_xref="GI:1066309"	Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS	TG1PAAISLSLFPQLEFGLSKLGRKYLIIITGNLWMPAFFIIFPGLQYNN LFTNSIVGIVGLFCFNAGPAVEAFIEKVSNNFERNFGRAMFCGVWALCASIVGI NLTINQVFWLGGSCALILAVLLFPKATDPASSATVANAGNHSFSLKLALELIR OPKLMFSLYIVGISTYDVDFQDFANFTSFATGEQGTFRVGVVITVMBELLNASIM FPAPLIIIRIGKKNALLAGTMSVRIIGSSFATSALSVWILKTLHMEVFPFLVGC KYITSOEVRPSATLYLVCFCFEKQLAMFMSVLNAGNYESIGQAVLVLGLVALGP TLISVFTSGGPGSLLRQNEVA" complement (20812. .23859) /gene="lacy2" complement (20812. .23859) /gene="lacy2" /note="EcoRI site normally present at the 3' end of lacZ has been removed by an uncharacterized mutation" /codon_start=1 /transl_table=11 /product="beta-galactosidase" /protein_id="AAC53652.1" /db_xref="GI:1066310" /translation="VVLRQWENPGVTQLNRLAAHPFASRNSEEARTRDPQQOLR SLNGEMRFAMPAPAPAEPSWLECDLPEADTVVPSNMOMHGYDAPITVNTYPI PPFVPTENTPQCYSLTFNVDSWLSQGTIRIFDGNVSAPHLNCGRWVGYGQDSRLP SEFDLSAFRAGENLAVMLVRLWSQSYLEDQMMRMSGIFRDLVSLHKPTQISDFH VATRPNDPFSRAVLBAEVMCGELRLYRLVTVSLWQGTQVAGSTAPFGGRIIDERG YADRVTLBLANVETKMSAEIPNLRYAVVELHTADGTLIEAACDVGPREVRIENGLL LNGKPLIRGVNRHREHPHQVMDQETMQDILLMKQNNFNNAVRCSHYNHPLWYT LCDRYGLVVDENAIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIWSL GNESGHAGVDEADLRYWIKSVDPSPVQYEGGADTATDIIICPMYARVDEDPFPAVP KWSIKWLSLPGTEYRPLILCEYAHAMGNSLGGFAKYQWAFRQYPRLOQGFVMDWQDS LKYDENPNPMSAGFDGDPNDPQFCWNGLVFADRTPHALTEAKHQOQFQFRLS GQTEVTSYVFRUSDNELLHMVALDKPLASGEVPLDVAPOCKQLJELPELPQES AQQLMTRVVRQPNATNSEAGHISAWQOURLAENLSVTLPAASHAIPLHILTSEMDPC IELGNRQWRFQSLQSMIGMKKQLTLPLRDQFTRAPLDNDIGVSEATRIDPNFAC VERKWAAGYQAEALLQCTADTLADAVLITTAHAWOQGTLFISRKYTRIDSGQM AITVDVEASDTDPHARIGLNCOLAQVAERNVWGLGPOENYPRDLTAACFDRWDPL SDMTYPPYFPPSENGURCTRELNYGPHQWGDQFQFNISRYSQOQLMETSHRLLHAE GTWLNIDGFHWGIGDDSWSPSVAEFQLSAGRHYQLVWCQK" complement (23875. .23958) /note="multiple cloning site" complement (23974. .24142) /note="rrnB1 terminator" 24476. .25666 /gene="tetr" 24476. .25666 /gene="tetr" /codon_start=1 /transl_table=11 /product="tetracycline resistance protein" /protein_id="AAC53653.1" /db_xref="GI:1066311"	QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTTCGGCTCAGTCATCGCCCAAGC 104 Db 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTTCGGCTCAGTCATCGCCCAAGC 3114 QY 105 TGGCGCTATCTGGGCATCGGGGAGGAAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 164 Db 3115 TGGCGCTATCTGGGCATCGGGGAGGAAAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 3174 QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAACGACG 224 Db 3175 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAACGACG 3234 QY 225 TTTACCATGATGATTCGGGAAGGTGTGCCCATGCACGCTTTTAAACGGTGAACTGTTCGTT 284 Db 3235 TTTACCATGATGATTCGGGAAGGTGTGCCCATGCACGCTTTTAAACGGTGAACTGTTCGTT 3294 QY 285 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGATTCGAGATGGTTCAGC 344 Db 3295 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGATTCGAGATGGTTCAGC 3354 QY 345 CCGAAGCGCATCAGCAACCCGCAACAATACCGCGCACAGCCGGAATCGCGTCCGGGTGTG 404 Db 3355 CCGAAGCGCATCAGCAACCCGCAACAATACCGCGCACAGCCGGAATCGCGTCCGGGTGTG 3414 QY 405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464 Db 3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474 QY 465 TGGATGCGCGAGAAATGGAACATGATATCCCGTGAATACCCCGTGAATACCCCGGCGCGCTCGTTC 524 Db 3475 TGGATGCGCGAGAAATGGAACATGATATCCCGTGAATACCCCGTGAATACCCCGGCGCGCTCGTTC 3534 QY 525 ATTCAGCTTTTGAACCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGCTTTTACAGC 584 Db 3535 ATTCAGCTTTTGAACCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGCTTTTACAGC 3594 QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAACACACGACGCT 631 Db 3595 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAACACACGACGCT 3641
misc_feature	RESULT 5 CVU39286 LOCUS DEFINITION Cloning vector TLF97-3, phage lambda lacZ translational fusion vector, complete sequence. ACCESSION U39286 VERSION U39286.1 KEYWORDS GI:1066312 SOURCE Cloning vector TLF97-3 ORGANISM Cloning vector TLF97-3 other sequences; artificial sequences; vectors. REFERENCE 1 (bases 1 to 42531) St Pierre,R. and Linn,T. A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ	CVU39286 42531 bp DNA linear SYN 13-APR-1996 Cloning vector TLF97-3, phage lambda lacZ translational fusion vector, complete sequence. U39286 U39286.1 GI:1066312 Cloning vector TLF97-3 Cloning vector TLF97-3 other sequences; artificial sequences; vectors. 1 (bases 1 to 42531) St Pierre,R. and Linn,T. A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ Gene 169 (1), 65-68 (1996) JOURNAL MEDLINE PUBMED 6186904 8635751 REFERENCE 2 (bases 1 to 42531) St Pierre,R. Direct Submission AUTHORS TITLE Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology, University of Western Ontario, London, Ontario N6A5C1, Canada JOURNAL FEATURES 1. 42531 /organism="Cloning vector TLF97-3" /mol_type="genomic DNA" /db_xref="taxon:43842" complement (19404. .19489) /note="rrnB2 terminator"
terminator	terminator	
gene	complement (35451 33603. .35451 /note="imm21 substitution"	
CDS	ORIGIN Query Match 82.8%; Score 585.4; DB 12; Length 42530; Best Local Similarity 99.8%; Pred. No. 2.5e-148;	

Query Match 82.8%; score 585.4; DB 12; Length 42531;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;

```
removed by an uncharacterized mutation"
5505
/note="BamHI site normally present at this site has been
removed by an uncharacterized mutation"
complement(19404..19489)
/note="rrnB2 terminator"
complement(19507..20760)
/gene="lacY"
complement(19507..20760)
/gene="lacY"
/codon_start=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53645.1"
/db_xref="GI:1051182"
/translation="MYVLYKNTNFMWGLPFFPYFIMGAYFPFPFIMLHDINHLSKSD
TGIFPAISLSFLPQPLGSLGDKIGURKYLWLIITGLVMFAFPFIFIPGLLQYN
ILVGSIVGIIYLGSCFNAGPAVEAFIEKVSRRNFEGFRAMFGCGVWALCASIVGI
MFTINQPFVYLGSCGICALILAVLLFFAKTDAPSSATVANAGVNHSAFSLKALELFR
QPKLWFLSLYVIGVSCIVDFDQDFANFTSFATGEOGTRVFGVVTMGELINASIM
FPALINIRIGKXALLAGTIMSVRIIGSSFATSALEWVILKTLHMFEPVFLVGCFF
KYITSQEVRSATILYVCFCEFFQLAMFMSVLGNMYESIGFQGYLVVLGLVALGF
TLISVTLGPGPGLLRQVNEVA"
complement(20812..23886)
/gene="lacZ"
complement(20812..23886)
/gene="lacZ"
complement(20812..23886)
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC53645.1"
/db_xref="GI:1051183"
/translation="MTMLTDSLAVLVQRDWNFCVLTQNLRLAHPFPFASWRNSEEAR
TRDPSQLRSLNGEMFAWFPAPPAEPSEWLECDLPEADTVVPSNMQMGYDAPIYT
NVITYIPNPPENTPGCYSLTFNVDSEWLGQGTRIIFDGVNSAFHLWCNCRWV
GYQDSRLPSSAFSAFAGENLAVMLRWSDGYLEDDMMWMSGIFRDSLLHK
PTTQISDPHATRFNDPSRAVLRAEQMCCGLDYLRTVLSLWQGETOVASGTAFCG
GEIIDRGGYADRTLRLNVPKLSAEIENLYRAVELHTAGTGLIEACADVGFR
EVLNGLLLGLLGLLVYVNRHEHFLHGVQWDEQTMVODIILMKQNNFNARVCRSH
YFNPLMYTLCDRYGLYVDEANITETHGMVPMRLTDPPRLWPAMSERVTRMVRQDRN
HPSVI1WSLGNESGHANDALYRWIKSVSPSPVQVEGGGATTAIDIICPMYARVD
EQDPPAPVKWSIKWLSLPGETRLPLICEYAHAMGNSLGGAFKYQWAFROYPRLGQ
FVMDVQSLIKYDENGPNWASGYGDFGTDNDQFCWNGVLVFADRTPHALTEAKHQ
QQFPQLSGQTIVTSEYLFHSDNELHWMVALDKPLASGEVPHLDVAPOGQKLE
LPFLQPBSAQTLVWVQPNATWSEAGHISAWQOQWLAENSLVTLPAASHAIEH
LTTSEMDFCIBLKNRWQFNRSQFLSQMTGDKKQLTLPLRDQFTRAPLNDIGVSE
ATRIDPNAWERWKAAGHYQAEALLQCTADTLADAVLITTAHAWHQGKTLFSRKT
YRIDSGQMALTVDEVASDTPHPARIGLNCQLAQVAVNMLGLGPOENVPDLRTAA
CFDRWDLPTDMYTPYPPSPENGLRCGTREINLPGHWRGDFQFNISRYSOQLMETS
HRHLHAEETWLNIDGPHMIGGDDSWSPSVAEFQLSAGRYHYQLWVCK"
complement(24049..24132)
/note="multiple cloning site"
complement(24148..24316)
/note="rrnB1 terminator"
24650..25840
/gene="tetR"
24650..25840
/gene="tetR"
/note="Description: tetracycline resistance gene;
tetracycline resistance protein"
/codon_start=1
/transl_table=11
/protein_id="AAC53647.1"
/db_xref="GI:1051184"
/translation="MKSNNALIVILGTVTLDAVGILVMPVLPGLLRDIVHSDSIASH
YGVLLAYALMOFLCAPVLGALSDFRGRPPVLLASLIGLATIDYAIMATTPVLTLYAG
RIVAGITATGACVAGAYADITDGEDRARHFGLMSACFGVGMVAGPVAGGLGLAISLH
APFLAAVLNALLGLGFLQMSHKGERRPMLRAENPVSSFRKAGMTIVALMTV
PFTQLGVQVPAALVIFGEDRFRWSATMIGLSLAVFGILHALAQAFVGTATKRFGE
KQAI1AGNAADALGVLLAFATGRWMAFFIMILLASGIGIMPALQOAMLSRDDDDHOG
QLOGSLAALTSLTISGTPLTVAIYASASTWNGLAWITVGAALYLVCLPALRRGAWSR
```

```
misc_feature
31747
ATST"
/note="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
33777..33625
/note="imm21 substitution"
39168
/note="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
ORIGIN
Query Match 82.8%; Score 585.4; DB 12; Length 42704;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
Db 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTCTTTTTCGGCTCAGTCATCGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGGAGAAAGCCGTCCTTTTCCCGCAGGTTGAAGCG 164
Db 3115 TGGCGCTATCTGGGCATCGGGGAGAAAGCCGTCCTTTTCCCGCAGGTTGAAGCG 3174
QY 165 GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAAAAACGACG 224
Db 3175 GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAAAAACGACG 3234
QY 225 TTTACCATGATGATTCGGGAAGGTGTGCCATGCACGCCCTTTTACCGTGACTCTTCGTT 284
Db 3235 TTTACCATGATGATTCGGGAAGGTGTGCCATGCACGCCCTTTTACCGTGACTCTTCGTT 3294
QY 285 CAGGCCACCTTGGGATACAGTTCGTCGGGCTTTTCGGGACACAGTTCCGGATCGTCAGC 344
Db 3295 CAGGCCACCTTGGGATACAGTTCGTCGGGCTTTTCGGGACACAGTTCCGGATCGTCAGC 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATCTGCCGTCCGTTG 404
Db 3355 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAATCTGCCGTCCGTTG 3414
QY 405 CAGATTAATGACAGCGGTGCGGCTGGGATATTAAGTCAGCAGGAGGAGGATTCCTGGC 464
Db 3415 CAGATTAATGACAGCGGTGCGGCTGGGATATTAAGTCAGCAGGAGGAGGATTCCTGGC 3474
QY 465 TGGATGCGCGAGAATGACATGATATACCCGTCAGTTACCCGCGCGCGCCCTCGTTTC 524
Db 3475 TGGATGCGCGAGAATGACATGATATACCCGTCAGTTACCCGCGCGCGCCCTCGTTTC 3534
QY 525 ATTACCGCTTTTGAACCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db 3535 ATTACCGCTTTTGAACCCGTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 3594
QY 585 GTGATGAGCAGCATGAAGATGCTCGACACGCTGCAGACACGACGACGACGCT 631
Db 3595 GTGATGAGCAGCATGAAGATGCTCGACACGCTGCAGACACGACGACGACGCT 3641
RESULT 7
LAMCQ 48502 bp DNA circular PHG 17-APR-2002
LOCUS Bacteriophage lambda, complete genome.
DEFINITION J02459 M17233 M24325 V00636 X00906
ACCESSION J02459.1 GI:215104
VERSION
KEYWORDS DNA-binding protein; circular; coat protein; complete genome;
origin of replication; repressor; unidentified reading frame.
SOURCE Bacteriophage lambda
ORGANISM Bacteriophage lambda
REFERENCE 1 (bases 1 to 12)
AUTHORS Wu, R. and Taylor, E.
TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide
sequence of the cohesive ends of bacteriophage lambda DNA
JOURNAL J. Mol. Biol. 57 (3), 491-511 (1971)
```


71209066 MEDLINE
4931680 PUBMED
REFERENCE
2 (bases 45493 to 45963)
AUTHORS
Imada, M. and Tsugita, A.
TITLE
Amino acid sequence of lambda phage endolysin
JOURNAL
Nature New Biol. 233, 230-231 (1971)
REFERENCE
3 (sites)
AUTHORS
Weigel, P.H., Englund, P.T., Murray, K. and Old, R.W.
TITLE
The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 70 (4), 1151-1155 (1973)
73215915 MEDLINE
4515813 PUBMED
REFERENCE
4 (bases 38597 to 38672)
AUTHORS
Dahlberg, J.E. and Blattner, F.R.
TITLE
In vitro transcription products of lambda DNA: Nucleotide sequences
and regulatory sites
JOURNAL
(in) Fox, C.F. and Robinson, W.S. (Eds.),
VIRUS RESEARCH. PROCEEDINGS OF 1973 ICH-UCLA SYMPOSIUM: 533-544;
Academic Press, New York (1973)
5 (bases 37945 to 38027)
AUTHORS
Maniatis, T., Ptashne, M., Backman, K., Kield, D., Flashman, S.,
Jeffrey, A. and Maurer, R.
TITLE
Recognition sequences of repressor and polymerase in the operators
of bacteriophage lambda
JOURNAL
Cell 5 (2), 109-113 (1975)
75185528 MEDLINE
1095210 PUBMED
REFERENCE
6 (bases 35583 to 35600)
AUTHORS
Kleid, D.G., Agarwal, K.L. and Khorana, H.G.
TITLE
The nucleotide sequence in the promoter region of the gene N in
bacteriophage lambda
JOURNAL
J. Biol. Chem. 250 (14), 5574-5582 (1975)
75189495 MEDLINE
167018 PUBMED
REFERENCE
7 (bases 35434 to 35618)
AUTHORS
Dahlberg, J.E. and Blattner, F.R.
TITLE
Sequence of the promoter-operator proximal region of the major
leftward RNA of bacteriophage lambda
JOURNAL
Nucleic Acids Res. 2 (9), 1441-1458 (1975)
76031664 MEDLINE
1178525 PUBMED
REFERENCE
8 (bases 37945 to 38018)
AUTHORS
Maniatis, T., Jeffrey, A. and Kleid, D.G.
TITLE
Nucleotide sequence of the rightward operator of phage lambda
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)
75158212 MEDLINE
1055375 PUBMED
REFERENCE
9 (bases 44588 to 44773)
AUTHORS
Sklar, J., Yot, P. and Weissman, S.M.
TITLE
Determination of genes, restriction sites, and DNA sequences
surrounding the 6S RNA template of bacteriophage lambda
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)
75217847 MEDLINE
1098044 PUBMED
REFERENCE
10 (bases 37905 to 37989)
AUTHORS
Walz, A., Pirrotta, V. and Ineichen, K.
TITLE
Lambda repressor regulates the switch between PR and P_{RM} promoters
JOURNAL
Nature 262 (5570), 665-669 (1976)
76267718 MEDLINE
958438 PUBMED
REFERENCE
11 (bases 37946 to 38039)
AUTHORS
Smith, G.R., Eisen, H., Reichardt, L. and Hedgepeth, J.
TITLE
Deletions of lambda phage locating a prm mutation within the
rightward operator
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)
76152323 MEDLINE
1062780 PUBMED
REFERENCE
12 (bases 35578 to 35667; 37903 to 38027)
AUTHORS
Ptashne, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R.,
Meyer, B. and Sauer, R.T.
TITLE
Autoregulation and function of a repressor in bacteriophage lambda
JOURNAL
Science 194 (4261), 156-161 (1976)
76271154 MEDLINE

959843 PUBMED
13 (bases 35578 to 35667)
AUTHORS
Humayun, Z., Jeffrey, A. and Ptashne, M.
TITLE
Completed DNA sequences and organization of repressor-binding sites
in the operators of phage lambda
JOURNAL
J. Mol. Biol. 112 (2), 265-277 (1977)
77209970 MEDLINE
875019 PUBMED
REFERENCE
14 (bases 38610 to 38732)
AUTHORS
Scherer, G., Hobom, G. and Kossel, H.
TITLE
DNA base sequence of the po promoter region of phage lambda
JOURNAL
Nature 265 (5590), 117-121 (1977)
77100320 MEDLINE
834253 PUBMED
REFERENCE
15 (bases 38041 to 38241)
AUTHORS
Roberts, T.M., Shimatake, H., Brady, C. and Rosenberg, M.
TITLE
Sequence of Cro gene of bacteriophage lambda
JOURNAL
Nature 270 (5634), 274-275 (1977)
78071724 MEDLINE
593347 PUBMED
REFERENCE
16 (bases 27616 to 28935)
AUTHORS
Davies, R.W., Schreier, P.H. and Buchel, D.E.
TITLE
Nucleotide sequence of the attachment site of coliphage lambda
JOURNAL
Nature 270 (5639), 757-760 (1977)
78071823 MEDLINE
593399 PUBMED
REFERENCE
17 (bases 37206 to 37263; 37914 to 37970)
AUTHORS
Humayun, Z.
TITLE
DNA sequence at the end of the CI gene in bacteriophage lambda
JOURNAL
Nucleic Acids Res. 4 (7), 2137-2143 (1977)
78011659 MEDLINE
909767 PUBMED
REFERENCE
18 (bases 27617 to 27934)
AUTHORS
Landy, A. and Ross, W.
TITLE
Viral integration and excision: structure of the lambda att sites
JOURNAL
Science 197 (4309), 1147-1160 (1977)
77258934 MEDLINE
331474 PUBMED
REFERENCE
19 (bases 39062 to 39170)
AUTHORS
Denniston-Thompson, K., Moore, D.D., Kruger, K.E., Furth, M.E. and
Blattner, F.R.
TITLE
Physical structure of the replication origin of bacteriophage
lambda
JOURNAL
Science 198 (4321), 1051-1056 (1977)
78054731 MEDLINE
929187 PUBMED
REFERENCE
20 (bases 44467 to 44807)
AUTHORS
Sklar, J.B.
TITLE
Structure and function of two regions of DNA controlling the
synthesis of prokaryotic RNAs
JOURNAL
Thesis (1977)
21 (sites)
AUTHORS
Adhya, S. and Gottesman, M.
TITLE
Control of transcription termination
JOURNAL
Annu. Rev. Biochem. 47, 967-996 (1978)
78234064 MEDLINE
354508 PUBMED
REFERENCE
22 (bases 13 to 72; 48391 to 48502)
AUTHORS
Nichols, B.P. and Donelson, J.E.
TITLE
178-Nucleotide sequence surrounding the cos site of bacteriophage
lambda DNA
JOURNAL
J. Virol. 26 (2), 429-434 (1978)
78197067 MEDLINE
666898 PUBMED
REFERENCE
23 (bases 37938 to 38016; 35589 to 35666)
AUTHORS
Flashman, S.M.
TITLE
Mutational analysis of the operators of bacteriophage lambda
JOURNAL
Mol. Gen. Genet. 166 (1), 61-73 (1978)
79114073 MEDLINE
368570 PUBMED
REFERENCE
24 (bases 37990 to 38982)
AUTHORS
Schwarz, E., Scherer, G., Hobom, G. and Kossel, H.
TITLE
Nucleotide sequence of cro, cII and part of the O gene in phage

Qy 345 CCGAAGCGCATCAGCAACCCGAAACATACCGCGCGACAGCCGGAACCTGCCGCGGTGTG 404
 Db 4021 CCGAAGCGCATCAGCAACCCGAAACATACCGCGCGACAGCCGGAACCTGCCGCGGTGTG 4080
 Qy 405 CAGATTATGACAGCGGTGCCGCGTGGGATATTACGTACGCGAGGACGGGTATCCTGGC 464
 Db 4081 CAGATTATGACAGCGGTGCCGCGTGGGATATTACGTACGCGAGGACGGGTATCCTGGC 4140
 Qy 465 TGGATGCCGAGAAATGACATGATATCCCGTGGATTACCGCGCGCGCGCGCTCGTTC 524
 Db 4141 TGGATGCCGAGAAATGACATGATATCCCGTGGATTACCGCGCGCGCGCGCTCGTTC 4200
 Qy 525 ATTACAGCTTTTGAACCCGCGGAGGACGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
 Db 4201 ATTACAGCTTTTGAACCCGCGGAGGACGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 4260
 Qy 585 GTGATGGAGCAGATGAAGATGCTGACACCGCTGCAAGACAC 625
 Db 4261 GTGATGGAGCAGATGAAGATGCTGACACCGCTGCAAGACAC 4301
 RESULT 10
 AF434926 6182 bp DNA circular SYN 04-APR-2002
 LOCUS Expression vector 409-MUT, complete sequence.
 DEFINITION AF434926
 ACCESSION AF434926
 VERSION AF434926.1 GI:16904161
 KEYWORDS Expression vector 409-MUT
 SOURCE Expression vector 409-MUT
 ORGANISM other sequences; artificial sequences; vectors.
 REFERENCE 1 (bases 1 to 6182)
 AUTHORS Zieler, H. and Huynh, C.O.
 TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
 JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
 MEDLINE 21830665
 PUBMED 11841506
 REFERENCE 2 (bases 1 to 6182)
 AUTHORS Zieler, H. and Huynh, C.O.
 TITLE Direct Submission
 JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
 FEATURES
 Location/Qualifiers
 1..6182
 /organism="Expression vector 409-MUT"
 /mol_type="genomic DNA"
 /db_xref="taxon:176968"
 /note="insect-specific"
 97..614
 /note="Autographa californica nuclear polyhedrosis virus hr5 enhancer"
 enhancer
 615..1201
 /note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
 promoter
 1266..1398
 /note="artificial intron with mutated 3' splice site"
 intron
 1541..3193
 /codon_start=1
 /product="firefly luciferase"
 /protein_id="AAL30784.1"
 /db_xref="GI:16904162"
 /translation="MEDAKNIKKGPAPPYPLEDGTAGCQLHKKRYALVPGTIAFTD AHIEVNTIYASYFMSVLAEMKRYGLNTHRIIVVCSSENSLQFMFVLGALFVGAV APANDIYNERELLSMNTISQPTVVFVSKGLQKLNQKLPFIQKILIMDSKTDYQG FSHMYFTIHPFGNEDYDFVPSFDRDKTIALIMNSSGTGLPKGVLPHTACVYR PQSHADPTFGNQIIPDITAILSVVPHFGFMFTLGLYICGFRVVLVYMFEEELFLRS LQDYKIQSALVPTLFSFPAKSTLIDKVDLSNLHEIASGAPLSKEVCEAVAKRFLHP GIRQGYGTEVTSAILITPEGDKPKGKVVVPEAKVVDLDTGKTLGVNORGLCV RGPIMMSYVNPENATLIDKDLHSGDIAWDEDEHFFIVDLKSLIKYKGYQVA PAELSELQLQHPNIFDAGVAGLPDDAGELPAVVVLEHGKMTKEIVDVVASQVTT AKLRGGVVVFDEVPKGLTGKLDARKIREILIKKGGKSKL"
 CDS
 misc_feature 3359..3464
 /note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"
 CDS 4381..5241
 /codon_start=1
 /product="beta-lactamase"
 /protein_id="AAL30785.1"
 /db_xref="GI:16904163"
 /translation="MSIQHPRVALIPFFAAFLPVPFAHPETLVKVKDAEDQLGARVGY IELDLSGKILSPRPRPEEPFMSTFKVLLCGAVLSRIDAGQEOELGRRIHYSONDLVE YSPVTEKHLTDGNTVRELCSAALTMSDNTAANLLLTITIGSPKELTAFLHNGGDVTRL DRWPELNEAIPNDRDITMPVAMATTLRLKLTGLLETLASRQGLTIIDWMEADKRVAGEL LRSLALPAGNFADKSGAGERGSRGIIAALGPDGKPSRIIVVIYTTGSOATMDERNRQIA EIGASLIKHW"
 rep_origin 5529..5948
 /note="ColE1"
 ORIGIN
 Query Match 81.3%; Score 574.6; DB 12; Length 6182;
 Best Local Similarity 99.3%; Pred. No. 1.8e-145;
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 45 ATCTAGCTGCATCGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
 Db 3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
 Qy 105 TGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGTCGCCCTTTTCCCGGAGGTTGAAGCG 164
 Db 3781 TGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGTCGCCCTTTTCCCGGAGGTTGAAGCG 3840
 Qy 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAAAAACGCAAG 224
 Db 3841 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAAAAACGCAAG 3900
 Qy 225 TTACCATGATGATTCCGGGAAGGTGCGCATGACGCCCTTTTAAAGGTGAACTGTTGCTT 284
 Db 3901 TTACCATGATGATTCCGGGAAGGTGCGCATGACGCCCTTTTAAAGGTGAACTGTTGCTT 3960
 Qy 285 CAGGCCACCTGGGATACAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
 Db 3961 CAGGCCACCTGGGATACAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020
 Qy 345 CCGAAGCGCATCAGCAACCCGAAACATACCGCGCGACAGCCGGAACCTGCCGCGGTGTG 404
 Db 4021 CCGAAGCGCATCAGCAACCCGAAACATACCGCGCGACAGCCGGAACCTGCCGCGGTGTG 4080
 Qy 405 CAGATTATGACAGCGGTGCCGCGCTGGGATATTACGTACGCGAGGACGGGTATCCTGGC 464
 Db 4081 CAGATTATGACAGCGGTGCCGCGCTGGGATATTACGTACGCGAGGACGGGTATCCTGGC 4140
 Qy 465 TGGATGCCGAGAAATGACATGATATCCCGTGGATTACCGCGCGCGCGCGCTCGTTC 524
 Db 4141 TGGATGCCGAGAAATGACATGATATCCCGTGGATTACCGCGCGCGCGCGCTCGTTC 4200
 Qy 525 ATTACAGCTTTTGAACCCGCGGAGGACGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
 Db 4201 ATTACAGCTTTTGAACCCGCGGAGGACGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 4260
 Qy 585 GTGATGGAGCAGATGAAGATGCTGACACCGCTGCAAGACAC 625
 Db 4261 GTGATGGAGCAGATGAAGATGCTGACACCGCTGCAAGACAC 4301
 RESULT 11
 AF434927 6182 bp DNA circular SYN 04-APR-2002
 LOCUS Expression vector 410-FOR, complete sequence.
 DEFINITION AF434927
 ACCESSION AF434927
 VERSION AF434927.1 GI:16904164
 KEYWORDS Expression vector 410-FOR
 SOURCE Expression vector 410-FOR
 ORGANISM other sequences; artificial sequences; vectors.

REFERENCE	1	(bases 1 to 6182)
AUTHORS	Zieler,H. and Huynh,C.Q.	
TITLE	Intron-dependent stimulation of marker gene expression in cultured insect cells	
JOURNAL	Insect Mol. Biol.	11 (1), 87-95 (2002)
MEDLINE	21830565	
PubMed	11841506	
REFERENCE	2	(bases 1 to 6182)
AUTHORS	Zieler,H. and Huynh,C.Q.	
TITLE	Direct Submission	
JOURNAL	Submitted (17-OCT-2001)	Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES	Location/Qualifiers	
source	1. .6182	
	/organism="Expression vector 410-FOR"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:176969"	
	/note="insect-specific"	
enhancer	97. .614	
	/note="Autographa californica nuclear polyhedrosis virus hr5 enhancer"	
promoter	615. .1201	
	/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"	
CDS	1328. .2980	
	/codon_start=1	
	/product="firefly luciferase"	
	/protein_id="AAL30786.1"	
	/db_xref="GI:16904165"	
	/translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKMKRYALVPGTIAFTVD AHLEVNITYAEYFEMSVRKAEAMKRYGLANTNHRIVVCSNSLQFFMPVLGALFIGAV APANDIYNRELLNSMNIQSPTVVFVSKGGLKILNVKKLP1QIKIIMDSKTDYQG FQSNYFTVSHLPGFNEYFVPESDRDKTIALIMNSGSGTGLPGVALPHRTACVR FSHARDIPGNQIIPDPAIILSVVPEFHGFCMTTGLVLCGPRVLMYRFEELFELRS LDYKIQSALLVTLSFFPAKSLIDUKYDLNLSHETASGAPLSKEVGEAVAKRFLHP GIKQGYATTSAILLTPGGDKPGAVGVFFFAKVVDDTGTGLVGNQKRGELCV RGPMSGYNNPEATNALDKDQGLHSGDIAYWDEHFFIVDRILKSLIKYKGYQVA PAELSLIQQHPIITFAGVAGLPDDAGELPAAVVLHGKWTMTEKEIYDVYASQVTT AKKLRGVGVFVDEVPKGLTGKLDARKIRELIKAKKGKSKL"	
intron	3196. .3328	
	/note="artificial intron"	
misc_feature	3359. .3464	
	/note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"	
CDS	4381. .5241	
	/codon_start=1	
	/product="beta-lactamase"	
	/protein_id="AAL30787.1"	
	/db_xref="GI:16904166"	
	/translation="MSTQHFVRVALIPFFAFLPFAHPETLVKVKDAEDQLGARVGY IELDPSVKILEFRRPERPPEPMSTPKVLLCGAVLSRI DAGQEQLGRRIHYSNDLVE YSPVTEKHALPDGMVTRELCSAALTMSDNTAANILLTTIGGKELTFLHMGDHLVL DRWPELNEALPNDERITPVMATLRLKLTIGELLTLASRQQLIDWHEADKVAQGLP LRSALPAGWFIADKSGAGERSGIIAALPGDKPSRIVVITYTTSQATMDERNRQLA EIGASLIKKW"	
rep_origin	5529. .5948	
	/note="ColE1"	
ORIGIN		
Query Match	81.3%	Score 574.6; DB 12; Length 6182;
Best Local Similarity	99.3%	Pred. No. 1.8e-145;
Matches	577; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db	3721	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
Qy	105	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCTTTTCCCGCAGGTTGAAGCG 164
Db	3781	TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCTTTTCCCGCAGGTTGAAGCG 3840
Qy	165	GCATGGAAAGAGTTTGGCGGAGATGACTGCTGCTGCATTGACGTTGAGCGAAACGCACG 224

Db	3841	GCATGGAAAGAGTTTGC	CGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAACGCACG	3900
Qy	225	TTTACCATGATGATTTCGGGAAGGTGTGGCCATGCACGCGCTTTAA	CGGTGAACCTGTTTCGTT	284
Db	3901	TTTACCATGATGATTTCGGGAAGGTGTGGCCATGCACGCGCTTTAA	CGGTGAACCTGTTTCGTT	3960
Qy	285	CAGGCCACCTGGGANTACAGTTTCGTGCGCGCTTTTCGGACACAGTTTCCGGATGGTCAGC	344	
Db	3961	CAGGCCACCTGGGANTACAGTTTCGTGCGCGCTTTTCGGACACAGTTTCCGGATGGTCAGC	4020	
Qy	345	CCGAAGCGCATCAGCAACCCGGAACAATACCGCGCGACAGCCGGAACATGCGCTGCGGTTG	404	
Db	4021	CCGAAGCGCATCAGCAACCCGGAACAATACCGCGCGACAGCCGGAACATGCGCTGCGGTTG	4080	
Qy	405	CAGATTAATGACAGCGGTGCGCGCTGGGATATTCGTTCAGCGAGGACGGGTATCCTGGC	464	
Db	4081	CAGATTAATGACAGCGGTGCGCGCTGGGATATTCGTTCAGCGAGGACGGGTATCCTGGC	4140	
Qy	465	TGGATGCCGAGAAATGACATGCATACCCGTGAGTTACCCGCGCGCGCGCTCGTTC	524	
Db	4141	TGGATGCCGAGAAATGACATGCATACCCGTGAGTTACCCGCGCGCGCGCTCGTTC	4200	
Qy	525	ATTACAGCTTTTGAACCCGTTGAGAGGAGCGGCACAGCTCGCGGTGCAATGTTTACAGC	584	
Db	4201	ATTACAGCTTTTGAACCCGTTGAGAGGAGCGGCACAGCTCGCGGTGCAATGTTTACAGC	4260	
Qy	585	GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAAC	625	
Db	4261	GTGATGAGCAGATGAAGATGCTCGACACGCTGCATGAAC	4301	
RESULT 12	AF434931	6182 bp	DNA	Circular SYN 04-APR-2002
LOCUS	AF434931	Expression vector 409-REV	complete sequence.	
DEFINITION	AF434931	AF434931.1	GI:16904176	
ACCESSION	AF434931	Expression vector 409-REV		
VERSION	AF434931.1	Other sequences; artificial sequences; vectors.		
KEYWORDS		1 (bases 1 to 6182)		
SOURCE		Zieler, H. and Huynh, C.Q.		
ORGANISM		Intron-dependent stimulation of marker gene expression in cultured insect cells		
REFERENCE		Insect Mol. Biol. 11 (1), 87-95 (2002)		
AUTHORS		21830665		
TITLE		2 (bases 1 to 6182)		
JOURNAL		Zieler, H. and Huynh, C.Q.		
MEDLINE		Direct Submission		
PUBMED		Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of		
AUTHORS		Parasitic Diseases, National Institutes of Health, 4 Center Drive		
TITLE		MSC 0425, Bethesda, MD 20892-0425, USA		
JOURNAL		Location/Qualifiers		
FEATURES		1. .6182		
source		/organism="Expression vector 409-REV"		
enhancer		/mol_type="genomic DNA"		
		/db_xref="taxon:176973"		
		/note="insect-specific"		
		97. .614		
		/note="Autographa californica nuclear polyhedrosis virus hr5 enhancer"		
promoter		615. .1201		
		/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"		
intron		complement(1236. .1368)		
		/note="artificial intron"		
CDS		1541. .3193		
		/codon_start=1		
		/product="firefly luciferase"		
		/protein_id="AAU30794.1"		

```
/db_xref="GI:16904177"
/translation="MEDAKNIKKGPAPFPYLEDGTAGBQLHKKMKRYALVPGTIAFTD
AHIEVNIYVYAEFMSVRLAEAMKRYGLNTNHRIVVCSNSLOQFMPVLGALFIGVAV
APANDIYNERELLSNMNISQPTVVFVSKGLQKILNVQKLPILQKIIIMDSKTDYQG
FQSMYTFVTHLPGFNEYDFVPSFRDKTIALIMNSSGTGAPKGVLPHPRTACVR
FSHARDPIFGNIIIPDTAILSVVPHFGFMFTLGLICGFRVLVAVRFEELFLRS
LQDYKIQSALVPTLFSFPAKSTLIDKYLNSLHEIASGGAPLSKEVGEAVAKRPHLP
GIQYQGLITETTSAILITPEGDKPGAVKVPVFEAKVVDLDTGKTLGVNQGELCV
RGPIMSGYVNNPEATNALIDKDWLHSGDIAVWDEDEHFFIVDRLSKSLIKYGYQVA
PAELISILLOHPNIFDAGVAGLPDDAGELPAAVVVLEHGKTMTEKEIVDVASQVTT
AKKLRGGVWFVDEVPKGLTGKLDARKIREILIKAKGGKSKL"
misc_feature
3359..3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
CDS
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30795.1"
/db_xref="GI:16904178"
/translation="MSIQHFRVALIPFFAARCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILSFPRPEERPMSTMTEFKVLLCGAVLSRIDAGQEOLGRRIHYSONDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNMGRHVTSL
DRWPELNEALPNDERDTTPVAVATTLRKLTGELLTLASRQQLIDWMEADKVKAGPL
LRSLALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSGQATMDERNRQIA
EIGASLTKHW"
rep_origin
5529..5948
/note="ColE1"
```

ORIGIN

```
Query Match      81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCACG 104
Db |||
Qy 3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCACG 3780
Db |||
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCTGCTTTTCCGGCAGGTTGAAGCG 164
Db |||
Qy 3781 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCTGCTTTTCCGGCAGGTTGAAGCG 3840
Db |||
Qy 165 GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGATGACGTTGAGCGGAAACGCGACG 224
Db |||
Qy 3841 GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGATGACGTTGAGCGGAAACGCGACG 3900
Db |||
Qy 225 TTTACCATGATGATTCCGGAAGGTGTGCCATGCACGCTTTTAAACGGTGAAGTTCGTT 284
Db |||
Qy 3901 TTTACCATGATGATTCCGGAAGGTGTGCCATGCACGCTTTTAAACGGTGAAGTTCGTT 3960
Db |||
Qy 285 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCGGGATGTCGACG 344
Db |||
Qy 3961 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCGGGATGTCGACG 4020
Db |||
Qy 345 CCGAAGCGCATCAGCAACCCGAAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTG 404
Db |||
Qy 4021 CCGAAGCGCATCAGCAACCCGAAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTG 4080
Db |||
Qy 405 CAGATTATGACAGCGGTGCGCGCTGGGATATTACGTCAGGAGGAGCGGTATCCTGGC 464
Db |||
Qy 4081 CAGATTATGACAGCGGTGCGCGCTGGGATATTACGTCAGGAGGAGCGGTATCCTGGC 4140
Db |||
Qy 465 TGGATGCCGAGAAATGACATGGATACCCCGTGAGTTACCCGGCGCGCGGCTCGTTC 524
Db |||
Qy 4141 TGGATGCCGAGAAATGACATGGATACCCCGTGAGTTACCCGGCGCGCGGCTCGTTC 4200
Db |||
Qy 525 ATTACAGCTTTTGAACCCCGTGAGAGACCGGCGAGACTCGCGGTGCAATGTGTTTACAGC 584
Db |||
Qy 4201 ATTACAGCTTTTGAACCCCGTGAGAGACCGGCGAGACTCGCGGTGCAATGTGTTTACAGC 4260
Db |||
Qy 585 GTGATGAGCAGATGAAGATCTGCACACGCTGCAGACAC 625
Db |||
Qy 4261 GTGATGAGCAGATGAAGATCTGCACACGCTGCATGAAAC 4301
Db |||
```

```
RESULT 13
AP434932
LOCUS AF434932 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 410-REV, complete sequence.
ACCESSION AF434932
VERSION AF434932.1 GI:16904179
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830865
PUBMED 11841506
REFERENCE
2 (bases 1 to 6182)
Zieler,H. and Huynh,C.Q.
Direct Submission
TITLE Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
JOURNAL
FEATURES
source
1..6182
/organism="Expression vector 410-REV"
/mol_type="genomic DNA"
/db_xref="taxon:176974"
/notes="insect-specific"
97..614
/notes="Autographa californica nuclear polyhedrosis virus hr5 enhancer"
615..1201
/notes="Autographa californica nuclear polyhedrosis virus immediate early promoter"
1328..2980
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30796.1"
/db_xref="GI:16904180"
/translation="MEDAKNIKKGPAPFPYLEDGTAGBQLHKKMKRYALVPGTIAFTD
AHIEVNIYVYAEFMSVRLAEAMKRYGLNTNHRIVVCSNSLOQFMPVLGALFIGVAV
APANDIYNERELLSNMNISQPTVVFVSKGLQKILNVQKLPILQKIIIMDSKTDYQG
FQSMYTFVTHLPGFNEYDFVPSFRDKTIALIMNSSGTGAPKGVLPHPRTACVR
FSHARDPIFGNIIIPDTAILSVVPHFGFMFTLGLICGFRVLVAVRFEELFLRS
LQDYKIQSALVPTLFSFPAKSTLIDKYLNSLHEIASGGAPLSKEVGEAVAKRPHLP
GIQYQGLITETTSAILITPEGDKPGAVKVPVFEAKVVDLDTGKTLGVNQGELCV
RGPIMSGYVNNPEATNALIDKDWLHSGDIAVWDEDEHFFIVDRLSKSLIKYGYQVA
PAELISILLOHPNIFDAGVAGLPDDAGELPAAVVVLEHGKTMTEKEIVDVASQVTT
AKKLRGGVWFVDEVPKGLTGKLDARKIREILIKAKGGKSKL"
complement(3166..3298)
/notes="artificial intron"
3359..3464
/notes="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30797.1"
/db_xref="GI:16904181"
/translation="MSIQHFRVALIPFFAARCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILSFPRPEERPMSTMTEFKVLLCGAVLSRIDAGQEOLGRRIHYSONDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNMGRHVTSL
DRWPELNEALPNDERDTTPVAVATTLRKLTGELLTLASRQQLIDWMEADKVKAGPL
LRSLALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSGQATMDERNRQIA
EIGASLTKHW"
5529..5948
/notes="ColE1"
rep_origin
ORIGIN
Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
```

Matches	577;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC	104						
Db	3721	ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC	3780						
Qy	105	TGGCGCTATCTGGGCATCGGGAGGAAGAACCGCGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC	164						
Db	3781	TGGCGCTATCTGGGCATCGGGAGGAAGAACCGCGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC	3840						
Qy	165	GCATGGAACAGTTTCGCGAGGATGACTGCTGCTGCAATGACCTTGAGCGCAAAACCCAGC	224						
Db	3841	GCATGGAACAGTTTCGCGAGGATGACTGCTGCTGCAATGACCTTGAGCGCAAAACCCAGC	3900						
Qy	225	TTTACCATGATGATTCGCGAAGGTGTGGCCATGCACGCTTTTAAACGCTGAACTGTTTCGTT	284						
Db	3901	TTTACCATGATGATTCGCGAAGGTGTGGCCATGCACGCTTTTAAACGCTGAACTGTTTCGTT	3960						
Qy	285	CAGGCCACTTGGGATACCATGTTCTGTCGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC	344						
Db	3961	CAGGCCACTTGGGATACCATGTTCTGTCGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC	4020						
Qy	345	CCGAAGCGCATCGCAACCGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTGTG	404						
Db	4021	CCGAAGCGCATCGCAACCGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTGTG	4080						
Qy	405	CAGATTAAATGACAGCGGTGCGGCTCGGATATTAACGTCAGCAGAGACGGGTATCTCTGGC	464						
Db	4081	CAGATTAAATGACAGCGGTGCGGCTCGGATATTAACGTCAGCAGAGACGGGTATCTCTGGC	4140						
Qy	465	TGATGCGCGAGAAATGGAATGATACCCCGTGAATACCCCGCGGCGCGCTCGTTC	524						
Db	4141	TGATGCGCGAGAAATGGAATGATACCCCGTGAATACCCCGCGGCGCGCTCGTTC	4200						
Qy	525	ATTACAGTTTTCGACCCGTGGAGAGCGGACACTCGCGTGCCTGCAATGTTTACAGC	584						
Db	4201	ATTACAGTTTTCGACCCGTGGAGAGCGGACACTCGCGTGCCTGCAATGTTTACAGC	4260						
Qy	585	GTATGAGCAGATGGAAGATGCTGCACACGCTGCAGAACAC	625						
Db	4261	GTATGAGCAGATGGAAGATGCTGCACACGCTGCAGAACAC	4301						

RESULT 14
AF434924
LOCUS AF434924 7421 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector pACTIN-LUC, complete sequence.
ACCESSION AF434924
VERSION AF434924.1 GI:16904155
KEYWORDS Expression vector pACTIN-LUC
SOURCE Expression vector pACTIN-LUC
ORGANISM Expression vector pACTIN-LUC
REFERENCE 1 (bases 1 to 7421)
Zieler, H. and Huynh, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 7421)
Zieler, H. and Huynh, C.Q.
TITLE Direct Substitution
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1. .7421
/organism="Expression vector pACTIN-LUC"
/mol_type="genomic DNA"
/db_xref="taxon:176966"
/note="insect-specific"
103. .2660
promoter

CDS	/note="Drosophila melanogaster actin5C promoter"
	2780. .4432
	/codon_start=1
	/product="firefly luciferase"
	/protein_id="AAL30780.1"
	/db_xref="GI:16904156"
	/translation="MEDAKNIKKGPAPFPYLEDGTAGEQIHKMKRYALVPGTIAFTD AHIEVNTIYAEYFEMSVLAEMKRYGLNTHRIIVVCSNSQJFMPVLGALFIGVAV AFANDIYNRELLNSMISQPTVVFVSKKGLKILNVQKLPFIQKIIIMDSKTDQG FOSMTYFVTHSLPFGNFYDVPESGRDKTIALIMNSSGSTGPKGVALPHRTACVR FSHARDPIGNQIIIPDAILSVVPHHGFMTTGLYLICGPRVLMYRPEBELFLR LDYKIQSALLVPTLSPFAKSTLDIKYDLSNLHEIASGAPLSKEVGEAVAKRHP GIRQGGTLTETTSAILITPEGDDKPGAVKVPFFFAKVVLDLTGKTLGVNQGELCV RPMMSGVNPNPEATNALIDKXMLHSDIAVWDEHFFIVDRUKSLIKYGYQVA PRAELISILQHPIEFADVAGLPDDDAGELPAAVVLEHGTMTKEIKVIDYVASQVTT AKKLGVVVFVDEVPKGLTGLDKARKIRILIKAKKGKSKL"
misc_feature	4598. .4703
	/note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"
	5620. .6480
CDS	/product="beta-lactamase"
	/codon_start=1
	/protein_id="AAL30781.1"
	/db_xref="GI:16904157"
	/translation="MSIQHFRVALIPFFPAFCPLPVFAHPETLVKVKDAEDQLGARVGY IEDLSNGIKLESFRPEERFPMWSTFKLLCGAVLSRIIDAGQQLGRRIRHYSQNDLVE YSPVTEKHLTDGMTRELCSAAITMSDNTAANLLTIIGPKELFTPLHMGDHLVRL DWPEELNEAIPIIDERDITMPVAMATTLRKLJLTGELLTLASRQQLIDMEADKVAQPL LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIIVYITTSQATMDERRQIA EIGASLIKHW"
rep_origin	6768. .7187
	/note="ColE1"
ORIGIN	
	Query Match 81.3%; Score 574.6; DB 12; Length 7421;
	Best Local Similarity 99.3%; Pred. No. 1.8e-145; Indels 0; Gaps 0;
	Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy	45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 104
Db	4960 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 5019
Qy	105 TGGCGCTATCTGGGCATCGGGAGGAAGAACCGCGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 164
Db	5020 TGGCGCTATCTGGGCATCGGGAGGAAGAACCGCGTCTTTTTCGGGTCTAGTCAATCGCCCAAGC 5079
Qy	165 GCATGGAACAGTTTCGCGAGGATGACTGCTGCTGCAATGACCTTGAGCGCAAAACCCAGC 224
Db	5080 GCATGGAACAGTTTCGCGAGGATGACTGCTGCTGCAATGACCTTGAGCGCAAAACCCAGC 5139
Qy	225 TTTACCATGATGATTCGCGAAGGTGTGGCCATGCACGCTTTTAAACGCTGAACTGTTTCGTT 284
Db	5140 TTTACCATGATGATTCGCGAAGGTGTGGCCATGCACGCTTTTAAACGCTGAACTGTTTCGTT 5199
Qy	285 CAGGCCACTTGGGATACCATGTTCTGTCGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC 344
Db	5200 CAGGCCACTTGGGATACCATGTTCTGTCGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC 5259
Qy	345 CCGAAGCGCATCGCAACCGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTGTG 404
Db	5260 CCGAAGCGCATCGCAACCGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTGTG 5319
Qy	405 CAGATTAAATGACAGCGGTGCGGCTCGGATATTAACGTCAGCAGAGACGGGTATCTCTGGC 464
Db	5320 CAGATTAAATGACAGCGGTGCGGCTCGGATATTAACGTCAGCAGAGACGGGTATCTCTGGC 5379
Qy	465 TGATGCGCGAGAAATGGAATGATACCCCGTGAATACCCCGCGGCGCGCTCGTTC 524
Db	5380 TGATGCGCGAGAAATGGAATGATACCCCGTGAATACCCCGCGGCGCGCTCGTTC 5439
Qy	525 ATTACAGTTTTCGACCCGTGGAGAGCGGACACTCGCGTGCCTGCAATGTTTACAGC 584
Db	5440 ATTACAGTTTTCGACCCGTGGAGAGCGGACACTCGCGTGCCTGCAATGTTTACAGC 5499

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 13:46:55 ; Search time 3474 Seconds
(without alignments)
1204.737 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatccatgcgtcaatttt.....aaaattgacgcattgggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	10	ACA55353 Eukaryoti
2	707	100.0	4613	10	ACA55360 Transform
C 3	705.4	99.8	8999	10	ACA55361 Transform
C 4	705.4	99.8	9009	10	ACA55362 Transform
C 5	705.4	99.8	9012	10	ACA55363 Transform
6	685	96.9	4943	10	ACA55369 Transform
7	685	96.9	4943	10	ACA55367 Transform
8	681.8	96.4	4944	10	ACA55368 Transform
9	585.4	82.8	48502	12	ADP70043 Nucleic a
10	580.6	82.1	4164	4	AAS44733 Bacteriop
C 11	566.2	80.1	26565	6	ABS78924 E. coli C
C 12	566.2	80.1	26565	10	ADH80491 Escherich
13	566.2	80.1	46819	9	ACD19059 E. coli 0
C 14	564.6	79.9	46897	10	ADCO0585 Enterohae
C 15	561.4	79.4	22306	9	ACD19238 E. coli 0
C 16	561.4	79.4	91740	10	ADCO0956 Enterohae
C 17	519.2	73.4	3661	10	ACA55354 Transform
C 18	511	72.3	4941	10	ACA55366 Transform
C 19	511	72.3	4952	10	ACA55365 Transform
C 20	499	70.6	4951	10	ACA55364 Transform

21	425.8	60.2	1427	6	ABQ50926	Abq50926 Oligonuc1
C 22	425.8	60.2	1427	6	ABQ50927	Abq50927 Oligonuc1
C 23	392.6	55.5	3822	5	AAS94485	Aas94485 DNA encod
C 24	382.6	54.1	1427	6	ABQ50928	Abq50928 Oligonuc1
C 25	382.6	54.1	1427	6	ABQ50929	Abq50929 Oligonuc1
C 26	223	31.5	34063	9	ACD19083	AcD19083 E. coli 0
C 27	223	31.5	44029	10	ADC00710	AdC00710 Enterohae
C 28	223	31.5	49650	10	ADC00365	AdC00365 Enterohae
C 29	222.2	31.4	134141	6	ABN83487	Abn83487 Escherich
C 30	171.2	24.2	822	5	AAS80174	Aas80174 DNA encod
C 31	96	13.6	847	2	Aa227249	Aa227249 Human sec
C 32	78	11.0	9980	10	ACA55352	ACA55352 pIAO-P/L
C 33	73.8	10.4	5194	4	AAC85595	Aac85595 Plasmid p
C 34	73.8	10.4	5194	9	ADA09845	Ada09845 PiggyBac
C 35	73.4	10.4	5194	4	AAC85596	Aac85596 Plasmid p
C 36	73.4	10.4	5194	9	ADA09846	Ada09846 PiggyBac
C 37	72.8	10.3	7670	10	ACA55351	ACA55351 pIAO-P/L
C 38	67.8	9.6	2480	8	ABV76195	Abv76195 T-richoplu
C 39	67.8	9.6	5679	4	AAC85602	Aac85602 Helper pl
C 40	67.8	9.6	5679	9	ADA09857	Ada09857 Helper pl
C 41	67.8	9.6	6723	4	AAC85593	Aac85593 psep70/op
C 42	67.8	9.6	6723	4	AAC85592	Aac85592 Plasmid p
C 43	67.8	9.6	6723	9	ADA09843	Ada09843 PiggyBac/
C 44	67.8	9.6	6723	9	ADA09842	Ada09842 Plasmid p
C 45	67.8	9.6	7560	4	AAC85599	Aac85599 Plasmid p

ALIGNMENTS

RESULT 1
ACA55353
ID ACA55353 standard; DNA; 707 BP.
AC ACA55353;
XX
AC ACA55353;
XX
DT 09-JUN-2003 (first entry)
XX
DE Eukaryotic transformation vector piggyBAC ITR cartridge.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; ITR cartridge.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
XX (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUA/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
XX Example 3; Fig 3(C1); 15ipp; English.
XX
XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents a minimal sequence cartridge of the eukaryotic transformation
CC vector piggyBac
XX
SQ Sequence 707 BP; 164 A; 174 C; 208 G; 161 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 707; DB 10; Length 707;
Best Local Similarity 100.0%; Pred. No. 2.6e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 1 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60

QY 61 ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 61 ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120

QY 121 TCGGGGAGGAAGACCGCGTCTTTTCCGGGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 121 TCGGGGAGGAAGACCGCGTCTTTTCCGGGAGGTTGAAGCGGCATGGAAGAGTTTG 180

QY 181 CCGAGGATGACTGCTGCTGCAATGAGCTTTGAGCGAAGAACGACGTTTACCATGATGATTC 240
Db 181 CCGAGGATGACTGCTGCTGCAATGAGCTTTGAGCGAAGAACGACGTTTACCATGATGATTC 240

QY 241 GGGAGGAGTGCCCATGACGCGCTTTAAACGGTGAACTGTTCTGTCAGGCCACCTGGGATA 300
Db 241 GGGAGGAGTGCCCATGACGCGCTTTAAACGGTGAACTGTTCTGTCAGGCCACCTGGGATA 300

QY 301 CCAGTTCTGCGGGCTTTTTCGGGACACAGTTCCCGATGCTGAGCCGAGCGCATCAGCA 360
Db 301 CCAGTTCTGCGGGCTTTTTCGGGACACAGTTCCCGATGCTGAGCCGAGCGCATCAGCA 360

QY 361 ACCCGAACAATACCGGCGACAGCGGAACTGCGTGGCGGTGTCAGATTAAATGACACGG 420
Db 361 ACCCGAACAATACCGGCGACAGCGGAACTGCGTGGCGGTGTCAGATTAAATGACACGG 420

QY 421 GTGCGGCGCTGGGATATTAACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGAGAAAT 480
Db 421 GTGCGGCGCTGGGATATTAACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGAGAAAT 480

QY 481 GGAATGATACCCCGTGAATTTACCGGCGGCGCGCTGTTCAATTCAGCTTTTGAAC 540
Db 481 GGAATGATACCCCGTGAATTTACCGGCGGCGCGCTGTTCAATTCAGCTTTTGAAC 540

QY 541 CCGTGGAGGACGGGACAGACTGCGGTGCAAAATGCTTTTACAGCGTGATGAGCAGATGA 600
Db 541 CCGTGGAGGACGGGACAGACTGCGGTGCAAAATGCTTTTACAGCGTGATGAGCAGATGA 600

QY 601 AGATGCTCGACACGCTGACAGAACACGACAGTAGATTAAACCCCTAGAAAGATTAATCATTTG 660
Db 601 AGATGCTCGACACGCTGACAGAACACGACAGTAGATTAAACCCCTAGAAAGATTAATCATTTG 660

QY 661 TGACGTAGCTTAAAGATTAATCATGCTGTAATTAATGACGATGGATGCC 707
Db 661 TGACGTAGCTTAAAGATTAATCATGCTGTAATTAATGACGATGGATGCC 707
```

RESULT 2
ACA55360

ID ACA55360 standard; DNA; 4613 BP.

XX AC ACA55360;

XX DT 06-JUN-2003 (first entry)

XX DE Transformation vector piggyBAC related plasmid pCRII-ITR.

KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
DR P-ESDB; ABU70357, ABU70358.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Disclosure; Fig 10(B); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
SQ Sequence 4613 BP; 1064 A; 1183 C; 1273 G; 1093 T; 0 U; 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 4613;
Best Local Similarity 100.0%; Pred. No. 6.7e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 294 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353

QY 61 ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db 354 ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 413

QY 121 TCGGGGAGGAAGACCGCGTCTTTTCCGGGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 414 TCGGGGAGGAAGACCGCGTCTTTTCCGGGAGGTTGAAGCGGCATGGAAGAGTTTG 473

QY 181 CCGAGGATGACTGCTGCTGCAATTTGACGCTTGAAGCGAAGAACGACGCTTTTACCATGATTC 240
Db 474 CCGAGGATGACTGCTGCTGCAATTTGACGCTTGAAGCGAAGAACGACGCTTTTACCATGATTC 533

QY 241 GGGAGGAGTGCGGCATGACGCGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
Db 534 GGGAGGAGTGCGGCATGACGCGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 593

QY 301 CCAGTTCTGCGGGCTTTTTCGGGACACAGTTCCGGATGGTCCAGCCCAAGCGCATCAGCA 360
Db 594 CCAGTTCTGCGGGCTTTTTCGGGACACAGTTCCGGATGGTCCAGCCCAAGCGCATCAGCA 653
```


Qy 361 ACCGAAACATACCGCGGACAGCGGAACTGCCGTGCGGTGTCAGATTAAATGACAGCG 420
 Db 654 ACCGAAACATACCGCGGACAGCGGAACTGCCGTGCGGTGTCAGATTAAATGACAGCG 713
 Qy 421 GTGGCGCGCTGGGATATTACGTACAGCGGAGGACGCGGTATCTGGCTGGATCCGCGAGAAAT 480
 Db 714 GTGGCGCGCTGGGATATTACGTACAGCGGAGGACGCGGTATCTGGCTGGATCCGCGAGAAAT 773
 Qy 481 GGACATGGATACCCCGGTGAGTTACCCGCGCGGCGCGCTCGTTCAATTCACGTTTTTGAAC 540
 Db 774 GGACATGGATACCCCGGTGAGTTACCCGCGCGGCGCGCTCGTTCAATTCACGTTTTTGAAC 833
 Qy 541 CCGTGGAGGACGGGACGACTCGCGGTGCAATGTTTACAGCGTGAATGAGCAGATGA 600
 Db 834 CCGTGGAGGACGGGACGACTCGCGGTGCAATGTTTACAGCGTGAATGAGCAGATGA 893
 Qy 601 AGATGCTCGACAGCTCGACACGAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
 Db 894 AGATGCTCGACAGCTCGACACGAGCAGCTAGATTAAACCTAGAAAGATAATCATATTG 953
 Qy 661 TGACGTACGTTAAAGATAATCATCGTAAATTAATGACGATGGGATCC 707
 Db 954 TGACGTACGTTAAAGATAATCATCGTAAATTAATGACGATGGGATCC 1000

RESULT 3

ACA55361/c
 ID ACA55361 standard; DNA; 8999 BP.
 XX AC ACA55361;
 XX DT
 XX 06-JUN-2003 (first entry)
 XX Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.
 XX PiggyBac; transposon; eukaryotic transformation vector; ds;
 KW transformed cell; transgenic; transgenic; plasmid; cyclic;
 KW circular.
 XX Synthetic.
 XX US2002173634-A1.
 XX 21-NOV-2002.
 XX 30-OCT-2001; 2001US-00001189.
 XX 31-OCT-2000; 2000US-0244677P.
 XX 01-NOV-2000; 2000US-0244984P.
 XX (FRAS/) FRASER M J.
 XX (LIXX/) LI X.
 XX (BEAM/) BEAM T.
 XX (HUAU/) HUA-VAN A.

Fraser MJ, Li X, Beam T, Hua-Van A;
 WPI; 2003-352597/33.

New DNA molecule in the transposon piggyBac, useful for transferring
 genes into host cells or embryos for transforming the cells of embryos
 that can be used in making transgenic organisms.

Example 11; Fig 12(B); 151pp; English.

The invention describes a DNA molecule comprising at least 163
 consecutive nucleotide base pairs of the 3' terminal region beginning at
 the 3' terminal base pair, and at least 125 consecutive nucleotide base
 pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 the piggyBac molecule. The region extends from the restriction site SacI
 to the end of the piggyBac molecule. The DNA molecule in the transposon
 piggyBac is useful for transferring genes into host cells or embryos for
 transforming the cells of embryos. The transformed cells or embryos are

CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac
 XX
 SQ Sequence 8999 BP; 2180 A; 2416 C; 2342 G; 2061 T; 0 U; 0 Other;
 Query Match 99.8%; Score 705.4; DB 10; Length 8999;
 Best Local Similarity 99.9%; Pred. No. 3e-212;
 Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GGATCCCATGCGCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
 Db 8950 GGATCCCATGCGCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8891
 Qy 61 ATCATATCTGCGGTCTTTTTCGCGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGCA 120
 Db 8890 ATCATATCTGCGGTCTTTTTCGCGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGCA 8831
 Qy 121 TCGGGAGGAGAAAGCCGCTGCTTTTCCGCGAGGTGGAAGCGGCATGGAAGAGTTG 180
 Db 8830 TCGGGAGGAGAAAGCCGCTGCTTTTCCGCGAGGTGGAAGCGGCATGGAAGAGTTG 8771
 Qy 181 CCGAGGATGACTGCTGCTGATTTGAGCGGAAACGACGTTTACCATGATGATTC 240
 Db 8770 CCGAGGATGACTGCTGCTGATTTGAGCGGAAACGACGTTTACCATGATGATTC 8711
 Qy 241 GGGAGGTGTGGCATGCGCCTTTAAACGCTGAACTGTTCTGTTACGCCACCTGGGATA 300
 Db 8710 GGGAGGTGTGGCATGCGCCTTTAAACGCTGAACTGTTCTGTTACGCCACCTGGGATA 8651
 Qy 301 CCAGTTGCTGCGGCTTTTCCGACACAGTTCCGGATGTTGTCAGCCCGGAGCGCATCAGCA 360
 Db 8650 CCAGTTGCTGCGGCTTTTCCGACACAGTTCCGGATGTTGTCAGCCCGGAGCGCATCAGCA 8591
 Qy 361 ACCGACAAATACGGGACGCGGAACTGCGGTGCGGTGTCAGATTAAATGACAGCG 420
 Db 8590 ACCGACAAATACGGGACGCGGAACTGCGGTGCGGTGTCAGATTAAATGACAGCG 8531
 Qy 421 GTGGCGCTGGGATATTACGTACAGGAGGACGGGTATCTGCTGGATCCGCGAGAAAT 480
 Db 8530 GTGGCGCTGGGATATTACGTACAGGAGGACGGGTATCTGCTGGATCCGCGAGAAAT 8471
 Qy 481 GGACATGGATACCCCGGTGAGTTACCCGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
 Db 8470 GGACATGGATACCCCGGTGAGTTACCCGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 8411
 Qy 541 CCGTGGAGGACGGGACGACTCGCGGTGCAATGTTTACAGCGTGAATGAGCAGATGA 600
 Db 8410 CCGTGGAGGACGGGACGACTCGCGGTGCAATGTTTACAGCGTGAATGAGCAGATGA 8351
 Qy 601 AGATGCTCGACAGCTGCGAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
 Db 8350 AGATGCTCGACAGCTGCGAGAACACGAGCTAGATTAAACCTAGAAAGATAATCATATTG 8291
 Qy 661 TGACGTACGTTAAAGATAATCATCGTAAATTAATGACGATGGGATCC 707
 Db 8290 TGACGTACGTTAAAGATAATCATCGTAAATTAATGACGATGGGATCC 8244

RESULT 4

ACA55362/c
 ID ACA55362 standard; DNA; 9009 BP.

XX AC ACA55362;
 XX DT
 XX 06-JUN-2003 (first entry)

Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.
 XX PiggyBac; transposon; eukaryotic transformation vector; ds;
 KW transformed cell; transgenic; transgenic; plasmid; cyclic;
 KW circular.

Db 1199 GGACATGGATACCCCGTGAGTTACCCGCGG-----CTCGTTCATTACGTTTTTGAAC 1252
 Qy 541 CCGTGGAGCAGCGGAGAGTCCGCGTGCAAAATGTGTTTTTACAGCGTGATGAGCAGATGA 600
 Db 1253 CCGTGGAGCAGCGGAGAGTCCGCGTGCAAAATGTGTTTTTACAGCGTGATGAGCAGATGA 1312
 Qy 601 AGATGCTCGACACGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAAGATTAATCATATTG 660
 Db 1313 AGATGCTCGACACGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAAGATTAATCATATTG 1372
 Qy 661 TGACGTAGCTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 707
 Db 1373 TGACGTAGCTTAAAGATAATCATGCGTAAATTTGACGCGATGGATCC 1419

RESULT 7

ACA55367
 ID ACA55367 standard; DNA; 4943 BP.

XX AC ACA55367;

XX XX 06-JUN-2003 (first entry)

XX DE Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.

XX KW PiggyBac; transposon; eukaryotic transformation vector; ds;
 XX KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 XX KW circular.

XX OS Synthetic.

XX PN US2002173634-A1.

XX PD 21-NOV-2002.

XX PF 30-OCT-2001; 2001US-00001189.

XX PR 31-OCT-2000; 2000US-0244677P.

XX PR 01-NOV-2000; 2000US-0244984P.

XX PA (FRAS/) FRASER M J.

XX PA (LIXX/) LI X.

XX PA (BEAM/) BEAM T.

XX PA (HUAU/) HUA-VAN A.

XX PI Fraser MJ, Li X, Beam T, Hua-Van A;

XX DR WPI; 2003-352597/33.

XX PS New DNA molecule in the transposon piggyBac, useful for transferring

XX PT genes into host cells or embryos for transforming the cells of embryos

XX PT that can be used in making transgenic organisms.

XX XX Example 6; Fig 18(B); 151pp; English.

XX CC The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac

XX SQ Sequence 4943 BP; 1221 A; 1305 C; 1245 G; 1172 T; 0 U; 0 Other;

Query Match 96.9%; Score 685; DB 10; Length 4943;
 Best Local Similarity 99.2%; Pred. No. 6.6e-206;
 Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATCGTCAATTTTACGAGACATATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
 Db 719 GGATCCCATCGTCAATTTTACGAGACATATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
 Qy 61 ATCATATCTGTCGGGTCTTTTTCGGCTCAGTATCGCCCAAGCTGCGCTATCTGGCA 120
 Db 779 ATCATATCTGTCGGGTCTTTTTCGGCTCAGTATCGCCCAAGCTGCGCTATCTGGCA 838
 Qy 121 TCGGGAGGAAGAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCTTGAAGAGAGTTTG 180
 Db 839 TCGGGAGGAAGAAGCCGCTCTTTTCCGCGAGGTTGAAGCGGCTTGAAGAGAGTTTG 898
 Qy 181 CCGAGGATGATCTGCTGCTCAATTGACGTTGAGCGAAACGACGTTTACCATGATTC 240
 Db 899 CCGAGGATGATCTGCTGCTCAATTGACGTTGAGCGAAACGACGTTTACCATGATTC 958
 Qy 241 GGGNAGGTGTGGCATGCACGCTTTTAAACGCTGAACCTGTTCTGTCAGGCCACCTGGGATA 300
 Db 959 GGGNAGGTGTGGCATGCACGCTTTTAAACGCTGAACCTGTTCTGTCAGGCCACCTGGGATA 1018
 Qy 301 CCAGTTCGTCGCGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
 Db 1019 CCAGTTCGTCGCGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 1078
 Qy 361 ACCGGAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTCAGATTAAATGACAGCG 420
 Db 1079 ACCGGAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGTGTCAGATTAAATGACAGCG 1138
 Qy 421 GTGGCGGCTGGGATATTACGTACGCGAGGACGCGGTATCTGCTGATGCCGAGAAAT 480
 Db 1139 GTGGCGGCTGGGATATTACGTACGCGAGGACGCGGTATCTGCTGATGCCGAGAAAT 1198
 Qy 481 GGACATGGATACCCCGTGAGTTTACCCGCGCGCGCTCGTTCATTACGTTTTCGAAC 540
 Db 1199 GGACATGGATACCCCGTGAGTTTACCCGCGCGCGCTCGTTCATTACGTTTTCGAAC 1252
 Qy 541 CCGTGGAGGACGCGGACAGACTCGCGGTGCAAAATGTGTTTTACGCGTATGAGCAGATGA 600
 Db 1253 CCGTGGAGGACGCGGACAGACTCGCGGTGCAAAATGTGTTTTACGCGTATGAGCAGATGA 1312
 Qy 601 AGATGCTCGACACGCTGCGAACAACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTG 660
 Db 1313 AGATGCTCGACACGCTGCGAACAACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTG 1372
 Qy 661 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGCGATGGATCC 707
 Db 1373 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGCGATGGATCC 1419

RESULT 8

ACA55368
 ID ACA55368 standard; DNA; 4944 BP.

XX AC ACA55368;

XX XX 06-JUN-2003 (first entry)

XX DE Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.

XX KW PiggyBac; transposon; eukaryotic transformation vector; ds;
 XX KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 XX KW circular.

XX OS Synthetic.

XX PN US2002173634-A1.

XX PD 21-NOV-2002.

XX PF 30-OCT-2001; 2001US-00001189.

XX PR 31-OCT-2000; 2000US-0244677P.

XX PR 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.
 PA (LIXX/) LI X.
 PA (BEAM/) BEAM T.
 PA (HUAV/) HUA-VAN A.
 XX Fraser MJ, Li X, Beam T, Hua-Van A;
 XX WPI; 2003-352597/33.
 XX
 XX New DNA molecule in the transposon piggyBac, useful for transferring
 PT genes into host cells or embryos for transforming the cells of embryos
 PT that can be used in making transgenic organisms.
 XX
 XX Example 6; Fig 19(B); 151pp; English.
 XX
 XX The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac
 XX
 XX Sequence 4944 BP; 1221 A; 1302 C; 1247 G; 1174 T; 0 U; 0 Other;
 SQ
 Query Match 96.4%; Score 681.8; DB 10; Length 4944;
 Best Local Similarity 99.9%; Pred. No. 6.8e-205;
 Matches 699; Conservative 0; Mismatches 2; Indels 6; Gaps 1;
 1 GGCATCCCATCGCTCAATTTTACCGAGACTATCTTCTAGGGTTAATCTAGCTCATCAGG 60
 719 GGCATCCCATCGCTCAATTTTACCGAGACTATCTTCTAGGGTTAATCTAGCTCATCAGG 778
 61 ATCATATCTGTCGGGTCTTTTTCGGCTCAGTATCCGCCAAGCTGCGCTATCTGGGCA 120
 779 ATCATATCTGTCGGGTCTTTTTCGGCTCAGTATCCGCCAAGCTGCGCTATCTGGGCA 838
 121 TCGGGGAGGAAGAACCGCTGCTTTTCCCGGAGGTGTTGAAGCGGCATGGAAGAGTTTG 180
 839 TCGGGGAGGAAGAACCGCTGCTTTTCCCGGAGGTGTTGAAGCGGCATGGAAGAGTTTG 898
 181 CCNAGGATGATCTGCTGCTATTCAGCTTGAGCGAAGAACCGCTTACCATGATGATTC 240
 899 CCGAGGATGATCTGCTGCTATTCAGCTTGAGCGAAGAACCGCTTACCATGATGATTC 958
 241 GGGAGGATGTCGCTGCTATTCAGCTTAAACGCTGAACTGTTTCCAGGCCACCTGGGATA 300
 959 GGGAGGATGTCGCTGCTATTCAGCTTAAACGCTGAACTGTTTCCAGGCCACCTGGGATA 1018
 301 CCAGTTCTGTCGGGTCTTTTCGGGACACAGTTTCCGGATGGTTCAGCCCGAAGCGCATCAGCA 360
 1019 CCAGTTCTGTCGGGTCTTTTCGGGACACAGTTTCCGGATGGTTCAGCCCGAAGCGCATCAGCA 1078
 361 ACCGACAAATACCGGCGACAGCCGGAACCTGCGGTGCGGTGTCAGATTAAATGACAGCG 420
 1079 ACCGACAAATACCGGCGACAGCCGGAACCTGCGGTGCGGTGTCAGATTAAATGACAGCG 1138
 421 GTGCGGCGCTGGGATATTTACGTGACGAGGACGGGTATCTTGGCTGGATGCGCGCAGAAAT 480
 1139 GTGCGGCGCTGGGATATTTACGTGACGAGGACGGGTATCTTGGCTGGATGCGCGCAGAAAT 1198
 481 GGACATGGATACCCCGTGAATTTACCCCGGCGCGCTCGTTCAATTCAGTTTTCGAAC 540
 1199 GGACATGGATACCCCGTGAATTTACCCCGGCGCGCTCGTTCAATTCAGTTTTCGAAC 1252
 541 CCGTGGAGGACGGGCGACATCTCGCGGTGCAATGTGTTTACAGCGTGTATGAGCAGATGA 600
 1253 CCGTGGAGGACGGGCGACATCTCGCGGTGCAATGTGTTTACAGCGTGTATGAGCAGATGA 1312

QY 601 AGATGCTCGACACGCTGCGAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 660
 DB 1313 AGATGCTCGACACGCTGCGAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
 QY 661 TGAAGCTACGTTAAAGATAATCATGCTGCTAAATTTGACGCGATGGATCC 707
 DB 1373 TGAAGCTACGTTAAAGATAATCATGCTGCTAAATTTGACGCGATGGATCC 1419
 RESULT 9
 ADP70043
 ID ADP70043 standard; DNA; 48502 BP.
 XX AC
 XX ADP70043;
 XX AC
 XX DT 26-AUG-2004 (first entry)
 XX DT
 XX DE Nucleic acid detection-related bacteriophage lambda DNA SeqID6.
 XX DE
 XX KW nucleic acid detection; nucleic acid amplification; microchip;
 XX KW isothermal; chip substrate; ds.
 XX OS Bacteriophage lambda.
 XX OS
 XX PN JP2004154008-A.
 XX PN
 XX PD 03-JUN-2004.
 XX PD
 XX PF 01-NOV-2002; 2002JP-00320330.
 XX PF
 XX PR 01-NOV-2002; 2002JP-00320330.
 XX PR
 XX PA (BIKE) EIKEN KAGAKU KK.
 XX PA
 XX DR WPI; 2004-445081/42.
 XX DR
 XX PT Detecting nucleic acid such as DNA, cDNA and RNA, involves performing
 PT nucleic acid amplification in microchip under isothermal conditions, and
 PT analyzing obtained amplified product on microchip.
 XX
 XX PS Example 2; SEQ ID NO 6; 49pp; Japanese.
 XX PS
 XX CC This invention relates to a novel method of detecting nucleic acid which
 CC comprises performing nucleic acid amplification in a microchip under
 CC isothermal conditions, and analysing the obtained amplified product on
 CC the microchip. The method is useful for detecting nucleic acid and
 CC enables performing a nucleic acid amplification reaction and its analysis
 CC continuously on the microchip without affecting the chip substrate. The
 CC method enables simple, reliable and cost-effective amplification,
 CC analysis and detection of nucleic acid in a microchip. The invention
 CC enables detection of trace amounts of nucleic acid with amplification
 CC efficiency. The present sequence is that of a bacteriophage lambda DNA
 CC sequence which was used in the exemplification of the invention.
 XX
 XX SQ Sequence 48502 BP; 12334 A; 11362 C; 12820 G; 11986 T; 0 U; 0 Other;
 Query Match 82.8%; Score 585.4; DB 12; Length 48502;
 Best Local Similarity 99.8%; Pred. No. 8e-174;
 Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCTCGCCCAAGC 104
 DB 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCTCGCCCAAGC 3114
 QY 105 TCGCGCTATCTGGGCATCGGGAGGAGAGCCGCTGCTTTTCCCGCGAGGTGAGAGCG 164
 DB 3115 TCGCGCTATCTGGGCATCGGGAGGAGAGCCGCTGCTTTTCCCGCGAGGTGAGAGCG 3174
 QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAGAACGACG 224
 DB 3175 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGAAGAACGACG 3234

QY 225 TTTACCATGATGATTCGCGGAAGGTGTGGCCATGCACGCGCTTTAAACGGTGAACCTGTTGCTT 284
Db TTTACCATGATGATTCGCGGAAGGTGTGGCCATGCACGCGCTTTAAACGGTGAACCTGTTGCTT 3294
QY 285 CAGGCCACTGGGATACCAAGTTCGTGCGGGCTTTTCCGGACACAGTTCGGATGGTGCAGC 344
Db CAGGCCACTGGGATACCAAGTTCGTGCGGGCTTTTCCGGACACAGTTCGGATGGTGCAGC 3354
QY 345 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGGTGCCGGTGTG 404
Db CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGGTGCCGGTGTG 3414
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCCTGGC 464
Db CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTTCAGCGAGGACGGGTATCCTGGC 3474
QY 465 TGGATGCCCGCAGAAATGACATGGATACCCCGTGAATACCGCGGGCGGGCGCTCGTTC 524
Db TGGATGCCCGCAGAAATGACATGGATACCCCGTGAATACCGCGGGCGGGCGCTCGTTC 3534
QY 525 ATTACAGTTTGTAAACCGGTGAGGACCGGACAGTTCGCGGTGCAAAATGTGTTTACAGC 584
Db ATTACAGTTTGTAAACCGGTGAGGACCGGACAGTTCGCGGTGCAAAATGTGTTTACAGC 3594
QY 585 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACACGCGAGCT 631
Db GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACACGCGAGCT 3641

RESULT 10

AAS44733

ID AAS44733 standard; DNA; 4164 BP.

AC AAS44733;

XX AAS44733;

DT 18-DEC-2001 (first entry)

Bacteriophage lambda full-length polynucleotide sequence #158.

KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
nervous system disorder; inflammatory disorder; cell differentiation; ds;
angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
genetic disorder; bone regeneration; tendon; ligament; tissue repair;
cytostatic; antirheumatic; antiarthritic; tendon; ligament; tissue repair;
antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
immunostimulant; analgesic; gene therapy.

OS Bacteriophage lambda.

XX WO200164834-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004926.

XX PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

PR 17-JUN-2000; 2000US-00597707.

PR 14-JUL-2000; 2000US-00616807.

PR 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

PI Drmanac R;

XX WPI, 2001-589862/66.

DR P-PSDB; AAU27833.

XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX Claim 1; SEQ ID NO 158; 153pp; English.

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
contig polynucleotides encoding polypeptides of the invention. The DNA
and protein sequences are useful for the treatment, diagnosis and
prevention of various types of disorder in a mammalian subject such as a
human, dog, monkey, mouse, hamster or rat. The disorders include cancers
such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
Wernicke disease, inflammatory disorders such as nephritis, Crohn's
disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
bowel disease. The sequences exhibit activity relating to angiogenesis,
cell proliferation, cell differentiation, stem cell growth factor,
activin or inhibin. Therefore, they can be used to manipulate stem cells
in culture to give rise to neuroepithelial cells that can be used to
augment or replace cells damaged by illness, accidental damage or genetic
disorders. The sequences may also be used for regeneration of bone,
cartilage, tendons and ligaments and in tissue repair and burn healing.
Note: Some sequences for this patent did not form part of the printed
specification, but were obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 4164 BP; 906 A; 1123 C; 1307 G; 828 T; 0 U; 0 Other;

Query Match 82.1%; Score 580.6; DB 4; Length 4164;
Best Local Similarity 99.3%; Pred. No. 7.8e-173;
Matches 583; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 237
QY 105 TGGCGCTATCTGGGCATCGGGGAGGAGAGCCGTCCTTTTCCCGAGGTTGAAGCG 164
Db 238 TGGCGCTATCTGGGCATCGGGGAGGAGAGCCGTCCTTTTCCCGAGGTTGAAGCG 297
QY 165 GCATGGAAGAGTTTTCGGGAGGATGACTGCTGCTGCATTCAGCTTGAGCGAAACGACG 224
Db 298 GCATGGAAGAGTTTTCGGGAGGATGACTGCTGCTGCATTCAGCTTGAGCGAAACGACG 357
QY 225 TTTACCATGATGATTCGGGAAGGTGTGCCATGCACGCGCTTTTAAACGGTGAACCTGCTT 284
Db 358 TTTACCATGATGATTCGGGAAGGTGTGCCATGCACGCGCTTTTAAACGGTGAACCTGCTT 417
QY 285 CAGGCCACCTGGGATACCACTTCGTGCGGGCTTTTCCGGACACAGTTCGGATCGTCAGC 344
Db 418 CAGGCCACCTGGGATACCACTTCGTGCGGGCTTTTCCGGACACAGTTCGGATCGTCAGC 477
QY 345 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGCGGTGTG 404
Db 478 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGCGGTGTG 537
QY 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTTACGTTCAGCGAGGACGGGTATCTTGGC 464
Db 538 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTTACGTTCAGCGAGGACGGGTATCTTGGC 597
QY 465 TGGATGCCCGCAGAAATGACATGGATACCCCGTGAATACCCGCGGGCGCGCTCGTTC 524
Db 598 TATATGCCGAGAAATGACATGGATACCCCGTGAATACCCGCGGGCGCGCTCGTTC 657
QY 525 ATTACAGCTTTTGAACCCGCGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
Db 658 ATTACAGCTTTTGAACCCGCGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 717
QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAAACACGCGAGCT 631


```
Db 718 GTGATGGAGCAGATAAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 764
RESULT 11
ABS78924/c
ID ABS78924 standard; DNA; 26565 BP.
XX
AC ABS78924;
XX
DT 17-DEC-2002 (first entry)
DE E. coli CFT073 genomic sequence #91.
XX
KW Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KW urinary tract infection; open reading frame; ORF; uropathogenic;
KW antibacterial; aropathic; nephrotropic; gene; ds.
XX
OS Escherichia coli.
XX
PN WO200259320-A2.
XX
PD 01-AUG-2002.
XX
PF 19-OCT-2001; 2001WO-US046833.
XX
PR 19-OCT-2000; 2000US-0242412P.
XX
PA (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI Blattner FR, Welch RA, Burland VD;
XX
DR WPI; 2002-691532/74.
XX
PT New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT useful for preventing or treating E. coli CFT073 infection in humans or
PT livestock.
XX
PS Claim 1; Page 276-291; 765pp; English.
XX
CC The present invention relates to polynucleotide sequences from the genome
CC of the pathogenic Escherichia coli strain CFT073. Almost all the
CC sequences present in E. coli CFT073 are absent in the previously
CC sequenced laboratory strain K-12. The polynucleotide sequences of the
CC invention are useful for preventing, diagnosing or treating E. coli
CC CFT073 infection in humans or livestock. The polynucleotide sequences are
CC useful for preventing urinary tract infections and pyelonephritis.
CC Likewise, the polypeptides encoded by the different open reading frames
CC (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC coli strains. ABS78834-ABS79085 represent genomic sequences from E. coli
CC strain CFT073
XX
SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;
Query Match 80.1%; Score 566.2; DB 6; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 45 ATCTAGCTGCATCAGGATCATATGCTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db ATCCAGCTGCATCAGGATCATATGCTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 3162
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 164
Db 3161 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCCGTGCCTTTTCCCGCAGGTTGAAGCG 3102
Qy 165 GCATGGAAGAGTTTTCGCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAGAACGCAAG 224
Db 3101 GCATGGAAGAGTTTTCGCGAGGACGACTGTTGCTGCAATTGACGTTGAGCGAAGAACGCAAG 3042
Qy 225 TTTCACCATGATGATTCGGGAGGTGTCGCCATGCACGCCCTTTAACGGTGAACGTTCGTT 284
Db 3041 TTACCATGATGATTCGGGAGGAGGTGTCGCCATGCATGCCCTTTAACGGTGAACGTTCGTT 2982
```

```
Qy 285 CAGGCCACCTGGGATACCAAGTTGCTCGCGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db 2981 CAGGCCACCTGGGATACCAAGTTGCTCGCGCTTTTCCGGACACAGTTCCGGATGGTCAGT 2922
Qy 345 CCGAAGCGCATCAGCAACCCGGAACATACCGGCGACAGCCGGAACCTGCCGTGCCGGTGTG 404
Db 2921 CCGAAGCGTATCAGCAACCCGGAACATACCGGCGACAGCCGGAACCTGCCGTGCCGGTGTG 2862
Qy 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACAGGAGACGGGATATCTCTGGC 464
Db 2861 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTACAGGAGACGGGATATCTCTGGC 2802
Qy 465 TGGATCCCGCAGAAATGGACATGATACCCCGTGAAGTACCCCGCGCGCGCGCTGCTTC 524
Db 2801 TGGATCCCGCAGAAATGGACATGATACCCCGTGAAGTACCCCGCGCGCGCGCTGCTTC 2742
Qy 525 ATTACAGTCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTTACAGC 584
Db 2741 ATTACAGTCTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTTACAGC 2682
Qy 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 631
Db 2681 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 2635
RESULT 12
ADH80491/c
ID ADH80491 standard; DNA; 26565 BP.
XX
AC ADH80491;
XX
DT 22-APR-2004 (first entry)
DE Escherichia coli CFT073 genome contig #91.
XX
KW ds; gene; Escherichia coli; CFT073; Escherichia coli CFT073 infection.
XX
OS Escherichia coli; CFT073.
XX
PN US2003165870-A1.
XX
PD 04-SEP-2003.
XX
PF 01-MAR-2002; 2002US-00085959.
XX
PR 01-MAR-2002; 2002US-00085959.
XX
PA (BLAT/) BLATTNER F R.
PA (WELC/) WELCH R A.
PA (BURL/) BURLAND V D.
XX
PI Blattner FR, Welch RA, Burland VD;
XX
DR WPI; 2003-863698/80.
XX
PT New nucleic acid of Escherichia coli CFT073, useful for preparing a
PT composition for diagnosing, treating or preventing infection caused by
PT Escherichia coli CFT073.
XX
PS Claim 1; SEQ ID NO 91; 4pp; English.
XX
CC The invention relates to an isolated Escherichia coli CFT073 nucleic acid
CC molecule. The nucleic acid is useful for preparing a composition for
CC diagnosing, treating or preventing infection caused by Escherichia coli
CC CFT073. The present sequence represents a contig of the Escherichia coli
CC CFT073 genome that is not present in Escherichia coli K-12.
XX
SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;
Query Match 80.1%; Score 566.2; DB 10; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```



```
Db 15275 ATTACGCTCTTTGAACCCGCTGGAGGACGGCGAGACCCGCGGTGCAAAATGTGTTTACAGC 15334
QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 631
Db 15335 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 15381

RESULT 14
ADC00585
ID ADC00585 standard; DNA; 46897 BP.
XX
AC ADC00585;
DT 04-DEC-2003 (first entry)
XX
DE Enterohaemorrhagic E. coli O157:H7-specific nucleic acid SEQ ID NO: 630.
XX ds; gene; enterohaemorrhagic; anti-bacterial.
XX Escherichia coli; O157:H7.
XX
OS JP2002355074-A.
PN
PD 10-DEC-2002.
XX
PF 24-JAN-2002; 2002JP-00015959.
XX
PR 24-JAN-2001; 2001JP-00112010.
XX
PA (UYTS-) UNIV TSUKUBA.
XX
XX WPI; 2003-451640/43.
DR
XX Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
PS Claim 2; SEQ ID NO 630; 2067pp; Japanese.
XX
XX The invention relates to a novel enterohaemorrhagic Escherichia coli
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E coli O157:H7 was determined. The present
CC sequence represents an E. coli O157:H7-specific nucleic acid of the
CC invention.
XX
SQ Sequence 46897 BP; 11872 A; 10948 C; 12731 G; 11346 T; 0 U; 0 Other;
Query Match 79.9%; Score 564.6; DB 10; Length 46897;
Best Local Similarity 97.6%; Pred. No. 3.1e-167;
Matches 573; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 45 ATCTAGTCGATCAGATCATATATCGTCGGGCTCTTTTCGGCTCAGTCATGCCCAAGC 104
Db 14795 ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTCGGCTCAGTCATGCCCAAGC 14854
QY 105 TGGCGCTATCTGGCATCGGGAGGAGAGCCGTCCTTTTCCCGGAGGTTGAAGC 164
Db 14855 TGGCGCTATCTGGCATCGGGAGGAGAGCCGTCCTTTTCCCGGAGGTTGAAGC 14914
QY 165 GCATGAAAGAGTTTGGCCGAGGATGACTGCTGCTGCAATTGACCTTGAGCGAAGAACGCACG 224
Db 14915 GCATGAAAGAGTTTGGCCGAGGATGACTGCTGCTGCAATTGACCTTGAGCGAAGAACGCACG 14974
QY 225 TTATCCATGATGATTCGGGAAGGTGGCCATGACGCCCTTTAAACGGTGAACCTGTCGTT 284
Db 14975 TTATCCATGATGATTCGGGAAGGTGGCCATGACGCCCTTTAAACGGTGAACCTGTCGTT 15034
QY 285 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCGGGACAGATTCGGGATGTCAGC 344
Db 15035 CAGGCCACCTGGGATACCCCGTCCCTCGCGACTGTCGGGACAGATTCGGGATGTCAGC 15094
```

RESULT 15

ACD19238/C

ID ACD19238 standard; DNA; 22306 BP.

XX

AC ACD19238;

XX

DT 27-OCT-2003 (revised)

DT 21-AUG-2003 (first entry)

XX

DE E. coli O157 unique DNA sequence OZID_251.

XX

KW OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;

KW food poisoning.

XX

OS Escherichia coli; strain O157:H7.

XX

PN US2003023075-A1.

XX

PD 30-JAN-2003.

XX

PF 01-APR-2002; 2002US-00114170.

XX

PR 04-DEC-1998; 98US-0110955P.

PR

03-DEC-1999; 99US-00453702.

XX

(BLAT/) BLATTNER F R.

PA (BURL/) BURLAND V D.

PA (PERN/) PERNA N T.

PA (PLUN/) PLUNKETT G.

PA (WELC/) WELCH R.

XX

PI Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;

XX

WPI; 2003-479497/45.

XX

New DNA sequences from Escherichia coli strain O157:H7, useful for

XX

Claim 16; SEQ ID NO 251; 33pp; English.

XX

The invention relates to an isolated DNA molecule comprising an E. coli strain O157:H7 sequence selected from a clostridial cytotoxin-like gene, a urease gene cluster, a RTX toxin-like gene cluster, a locus of enterocyte effacement and 2 genes from its associated lymphocytic phage 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli O157:H7 can cause food poisoning, specifically acute haemorrhagic colitis (which can develop into haemolytic uraemic syndrome). Also included are an isolated DNA molecule comprising a nucleotide sequence identical to at least 25 contiguous nucleotides contained in DNA sequences selected from ACD18988-ACD19242 (being 255 E.coli O157 DNA sequences which are not

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 14:11:50 ; Search time 27480 Seconds
(without alignments)
979.310 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggaatccatgcgtcaatttt.....aaaattgacgtggatccc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	585.4	82.8	682	7	CK781302
C 2	585.4	82.8	703	7	CF537771
C 3	585.4	82.8	711	7	CF743678
C 4	585.4	82.8	853	9	CF65566
C 5	584.4	82.7	703	5	BQ154655
C 6	584.4	82.7	719	6	CD350897
C 7	583.8	82.6	751	9	CR087413
C 8	580.4	82.1	716	9	CR131675
C 9	580.4	82.1	786	9	CR108810
C 10	565.4	80.0	691	6	CD350776
C 11	564	79.8	762	9	CR035207
C 12	563.4	79.7	749	6	CD351273
C 13	562	79.5	733	9	CR139473
C 14	555	78.5	633	9	CR167678
C 15	554.4	78.4	791	6	CB520716
C 16	550.4	77.9	630	9	CR013093
C 17	549	77.7	702	9	CR084987
C 18	545.4	77.1	611	4	B1423083
C 19	543.4	76.9	716	7	CK781284
C 20	540.8	76.5	716	7	CF851544
C 21	537.2	76.0	706	1	AV731514
C 22	535.4	75.7	730	9	CR077673
C 23	525	74.3	635	9	EX982794
C 24	521.8	73.8	536	9	CR026633

ALIGNMENTS

RESULT 1
CK781302/c
LOCUS CK781302 682 bp mRNA linear EST 23-FEB-2004
DEFINITION UI-M-GIO-clp-g-05-0-UI.r1 NIH_BMAP_GIO Mus musculus CDNA clone
IMAGE:30619060 5', mRNA sequence.

ACCESSION CK781302
VERSION CK781302.1 GI:42746980
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
(BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. .682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619060"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (tr. phage resistant)"
/clone_lib="NIH_BMAP_GIO"

FEATURES
source
notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag

D**b**

24 TGGCGCTATCTGGGCATCGGGAGGAAGAGC

REFERENCE 1 (bases 1 to 719)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: csapbe@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers
 1..719
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6853243"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_hosts="DH10B (tr1 phage resistant)"
 /clone_lib="NIH BMAP G10"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaudo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGCAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 82.7%; Score 584.4; DB 6; Length 719;
 Best Local Similarity 99.7%; Pred. No. 4.8e-164;
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
 DB 663 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 604

QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTTGAAGCG 164
 DB 603 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTTGAAGCG 544

QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAGAAACGCAAG 224
 DB 543 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAGAAACGCAAG 484

QY 225 TTACCATGATGATTCCGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACCTGTTGTT 284
 DB 483 TTACCATGATGATTCCGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACCTGTTGTT 424

QY 285 CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTTCCGGATGTTGTCAGC 344
 DB 423 CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTTCCGGATGTTGTCAGC 364

QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACACGCGGAACTCGCCGTGCCGGTGTG 404
 DB 363 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACACGCGGAACTCGCCGTGCCGGTGTG 304

QY 405 CAGATTAAATGACACCGGTGCGGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 464

DB 303 CAGATTAAATGACACCGGTGCGGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 244
 QY 465 TGGATGCCGAGAAATGGACATGATACCCCGTAGTATACCCGCGGCGCGCTCGTTC 524
 DB 243 TGGATGCCGAGAAATGGACATGATACCCCGTAGTATACCCGCGGCGCGCTCGTTC 184
 QY 525 ATTACAGTTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGC 584
 DB 183 ATTACAGTTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTACAGC 124
 QY 585 GTGATGAGCAGATGAAGATGCTCGACAGCTCAGAAACACGAGCT 631
 DB 123 GTGATGAGCAGATGAAGATGCTCGACAGCTCGAGAACACGAGCT 77

RESULT 7

CR087413/c

LOCUS 751 bp DNA linear GSS 05-JUL-2004
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHPP438f24, genomic survey sequence.

CR087413

ACCESSION CR087413.1 GI:49821005

VERSION GSS: genome survey sequence; MICR.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 751)

AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,

Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,

Rogers,J. and Bradley,A.

Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>

FEATURES

Location/Qualifiers

1..751

source

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHPP438f24"

/clone_lib="MHPP"

ORIGIN

Query Match 82.6%; Score 583.8; DB 9; Length 751;
 Best Local Similarity 99.7%; Pred. No. 7.3e-164;
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
 DB 623 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 564

QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTTGAAGCG 164
 DB 563 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCGTCCCTTTTCCCGGAGGTTGAAGCG 504

QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAGAAACGCAAG 224
 DB 503 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAGAAACGCAAG 444

QY 225 TTACCATGATGATTCCGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACCTGTTGTT 284
 DB 443 TTACCATGATGATTCCGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACCTGTTGTT 384

QY 285 CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTTCCGGATGTTGTCAGC 344
 DB 383 CAGGCCACCTGGGATACAGTTCTGTCGGGCTTTTCCGGACACAGTTTCCGGATGTTGTCAGC 324

QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACACGCGGAACTCGCCGTGCCGGTGTG 404
 DB 323 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACACGCGGAACTCGCCGTGCCGGTGTG 264

QY 405 CAGATTAAATGACACCGGTGCGGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 464

ORIGIN	Program Coordinator
Query Match	80.0%; Score 565.4; DB 6; Length 691;
Best Local Similarity	99.8%; Pred. No. 2.4e-158;
Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0	

ORIGIN	Query Match	79.8%	Score 564	DB 9	Length 762
	Best Local Similarity	100.0%	Pred. No. 6.5e-158		
	Matches 564	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY 68 CGTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGGCTATCTGGGCATCGGGGA 127
DB 2 CGTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGGCTATCTGGGCATCGGGGA 61
QY 128 GGAAGAAGCCGCTGCTTTTTCGGAGAGTGTGAAGCGGCATGGAAGAGTTTGGCAGGA 187
DB 62 GGAAGAAGCCGCTGCTTTTTCGGAGAGTGTGAAGCGGCATGGAAGAGTTTGGCAGGA 121
QY 188 TGACTGCTGCTGCAATGAGTGTGAGCAAAAGCAGCTTTACCATGATGATTCGGGAGG 247
DB 122 TGACTGCTGCTGCAATGAGTGTGAGCAAAAGCAGCTTTACCATGATGATTCGGGAGG 181
QY 248 TGTGGCCATGACAGCGCTTTTAAACGCTGAACTGTTCCTTCAGGCGCACTTGGGATACCAATTC 307
DB 182 TGTGGCCATGACAGCGCTTTTAAACGCTGAACTGTTCCTTCAGGCGCACTTGGGATACCAATTC 241
QY 308 CTCGGGCTTTTTCGGGACACATGTCGGATGTCAGCCCGAAGCGCATCAGCAACCCGAA 367
DB 242 CTCGGGCTTTTTCGGGACACATGTCGGATGTCAGCCCGAAGCGCATCAGCAACCCGAA 301
QY 368 CAATACCGGCGACAGCGGAACTGCGGTGCGGTGTGCAGATTAAATGACAGGGTGGCGC 427
DB 302 CAATACCGGCGACAGCGGAACTGCGGTGCGGTGTGCAGATTAAATGACAGGGTGGCGC 361
QY 428 GCTGGGATATTACGTCAGCGAGGACGGTATCTCTGGCTGGATGCCGCAAGAAATGGACATG 487
DB 362 GCTGGGATATTACGTCAGCGAGGACGGTATCTCTGGCTGGATGCCGCAAGAAATGGACATG 421
QY 488 GATACCCCGTGAGTTACCGCGGCGGCGCTGCTGTTCAATTCAGCTTTTGAACCCGTGGA 547
DB 422 GATACCCCGTGAGTTACCGCGGCGGCGCTGCTGTTCAATTCAGCTTTTGAACCCGTGGA 481
QY 548 GGACGGGCGAGTCCGGTGCAATGTCTTTTACAGCGTGATGGAGCAGATGCAAGTCT 607
DB 482 GGACGGGCGAGTCCGGTGCAATGTCTTTTACAGCGTGATGGAGCAGATGCAAGTCT 541
QY 608 CGACACGCTGCGAACAACGCGAGCT 631
DB 542 CGACACGCTGCGAACAACGCGAGCT 565
```

```
RESULT 12
CD351273/c
LOCUS UI-M-G10-csh-f-22-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:6853367 5', mRNA sequence.
ACCESSION CD351273.1 GI:31142608
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contract: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousef.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
```

```
FEATURES
source
```

```
Seq primer: pYX-5,
Location/Qualifiers
1..749
/organism="Mus musculus"
/mol_type="mRNA"
```

```
RESULT 13
CD351273/c
LOCUS
```

```
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
```

```
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IWAGE:6853367"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGCAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

ORIGIN

```
Query Match 79.7%; Score 563.4; DB 6; Length 749;
Best Local Similarity 99.8%; Pred. No. 9.8e-158;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGATCAGGATCATATCGTCGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 565 ATCCAGCTGTCATCAGGATCATATCGTCGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 506
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCTTTTCCCGGAGTTGAAGCG 164
DB 505 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCTTTTCCCGGAGTTGAAGCG 446
QY 165 GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCAATGACGTTGAGCGGAAACGACG 224
DB 445 GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCAATGACGTTGAGCGGAAACGACG 386
QY 225 TTTACCATGATGATTCGGGAGGTGTGCCATGACGCGCTTTAACGCTGACCTTCGTT 284
DB 385 TTTACCATGATGATTCGGGAGGTGTGCCATGACGCGCTTTAACGCTGACCTTCGTT 326
QY 285 CAGGCCACCTGGGATACAGTTTCGTCGGGCTTTTCGGGACACAGTTCCGGATGCTCAGC 344
DB 325 CAGGCCACCTGGGATACAGTTTCGTCGGGCTTTTCGGGACACAGTTCCGGATGCTCAGC 266
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGGCGACAGCCGGAACCTGCGGTCGTG 404
DB 265 CCGAAGCGCATCAGCAACCCGAAACATACCGGGCGACAGCCGGAACCTGCGGTCGTG 206
QY 405 CAGATTAATGACAGCGGTGGCGCTGGGATATTAATCTCAGCGAGGAGGGTATCTCTGGC 464
DB 205 CAGATTAATGACAGCGGTGGCGCTGGGATATTAATCTCAGCGAGGAGGGTATCTCTGGC 146
QY 465 TGGATGCGGCAAGAAATGGACATGGATACCCCGTGAATACCCGCGGGCGGCTCGTTC 524
DB 145 TGGATGCGGCAAGAAATGGACATGGATACCCCGTGAATACCCGCGGGCGGCTCGTTC 86
QY 525 ATTACGCTTTTGAACCCGCTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
DB 85 ATTACGCTTTTGAACCCGCTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 26
QY 585 GTGATGGAGCAGATGAAGATGCTCG 609
DB 25 GTGATGGAGCAGATGAAGATGCTCG 1
```

chromosome engineering clone MHP274b23, genomic survey sequence.	
ACCESSION	CR139473
VERSION	CR139473.1
KEYWORDS	GSS; genome survey sequence; MICBR.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 733)
AUTHORS	Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,D., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
TITLE	Direct Submision
JOURNAL	Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICBR
FEATURES	Location/Qualifiers
source	1..733
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/clone="MHP274b23"
	/clone_lib="MHP"
ORIGIN	
Query Match	79.5%; Score 562; DB 9; Length 733;
Best Local Similarity	99.3%; Pred. No. 2.6e-157;
Matches	573; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy	56 TCAGGATCATATCGTCGGGCTCTTTTTCGGCTCAGTCATCGCCAAAGCTGGCGCTATCT 115
Db	733 TCAGGATCATATCGTCGGGCTCTTTTTCGGCTCAGTCATCGCCAAAGCTGGCGCTATCT 674
Qy	116 GGGCATCGGGAGGAAGAGCCGTCGCTTTTCCGGGAGGTTGAAGGGCATGGAAAGA 175
Db	673 GGGCATCGGGAGGAAGAGCCGTCGCTTTTCCGGGAGGTTGAAGCGGCATGGAAAGA 614
Qy	176 GTTTGCCGAGATGACTGCTCTGCATTGACGTTTGAGCGAAACGACGTTTACCATGAT 235
Db	613 GTTTGCCGAGATGACTGCTCTGCATTGACGTTTGAGCGAAACGACGTTTACCATGAT 554
Qy	236 GATTCCGGAGGTTGGCCATGCACGCTTTTAAACGGTGAACTGTTGTTCAAGGCCACCTG 295
Db	553 GATTCCGGAGGTTGGCCATGCACGCTTTTAAACGGTGAACTGTTGTTCAAGGCCACCTG 494
Qy	296 GGATACCAATTCGTCGGGCTTTTCCGACACAGTTCCGATGTCAGCCCGAGACCGCAT 355
Db	493 GGATACCAATTCGTCGGGCTTTTCCGACACAGTTCCGATGTCAGCCCGAGACCGCAT 434
Qy	356 CAGCAACCCGAAACAATACCGCGCACAGCCGGAACGTCGG - TGCCGGTGTGCAGATTAATG 414
Db	433 CAGCAACCCGAAACAATACCGCGCACAGCCGGAACGTCGNNCTGCCNGTGTGCAGATTAATG 374
Qy	415 ACAGCGTGCAGCGCTGGGATATTACGTCACGAGGACGGGTATCTCTGGCTGGATGCGCG 474
Db	373 ACAGCGTGCAGCGCTGGGATATTACGTCACGAGGACGGGTATCTCTGGCTGGATGCGCG 314
Qy	475 AGAATGGACATGGATACCCCGTGAATTTACCGGGGGGCGGCTCGTTTCATTACAGTTT 534
Db	313 AGAATGGACATGGATACCCCGTGAATTTACCGGGGGGCGGCTCGTTTCATTACAGTTT 254
Qy	535 TTGAACCCGTCGAGGACGGGCAGACTCCGCGGTGCAAAATGTGTTTTACACGCTGATGGAGC 594
Db	253 TTGAACCCGTCGAGGACGGGCAGACTCCGCGGTGCAAAATGTGTTTTACACGCTGATGGAGC 194
Qy	595 AGATGAAGATGCTGCACACGCTGCAGAAACACGCGAGCT 631
Db	193 AGATGAAGATGCTGCACACGCTGCAGAAACACGCGAGCT 157

RESULT 14	CR167678/c	LOCUS	CR167678	633 bp	DNA	linear	GSS 06-JUN-2004
DEFINITION			Reverse strand read from insert in 3'HPRT insertion targeting and				

ACCESSION	Chromosome engineering clone	MHP161k02	genomic survey sequence.
VERSION	CR167678		
KEYWORDS	CR167678.1	GI:49946527	
SOURCE	GSS; genome survey sequence; M1CER.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 633)		
TITLE	Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L., Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y., Rogers, J. and Bradley, A.		
JOURNAL	Direct Submission		
FEATURES	Submitted (20-PEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/M1CER		
source	Location/Qualifiers		
ORIGIN	1. 633		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
	/clone="MHP161k02"		
	/clone_lib="MHP"		
Query Match	78.5%;	Score 555;	DB 9; Length 633;
Best Local Similarity	100.0%;	Pred. No. 3.2e-155;	
Matches	555; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
Qy	77	TTTTTTCGGCTCAGTCATCCGCAAGCTGGCGTATCTGGGCATCTGGGAGGAGGAAGC	136
Db	633	TTTTTTCGGCTCAGTCATCCGCAAGCTGGCGTATCTGGGCATCTGGGAGGAGGAAGC	574
Qy	137	CCGTGCCCTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTTCGCGAGATGACTGCTG	196
Db	573	CCGTGCCCTTTCCGCGAGGTTGAAGCGGCATGGAAGAGTTTTCGCGAGATGACTGCTG	514
Qy	197	CTGCATTGACGTTTGAGCGAAACGACGTTTACCATGATGATTCGGGAAGGTGTGGCCAT	256
Db	513	CTGCATTGACGTTTGAGCGAAACGACGTTTACCATGATGATTCGGGAAGGTGTGGCCAT	454
Qy	257	GCAGCCTTTAAACGGTGAACGTGTTTCAGTTCAGGCCACCTGGGATACGATTCGTCGGGCT	316
Db	453	GCAGCCTTTTTAAACGGTGAACGTGTTTCAGTTCAGGCCACCTGGGATACGATTCGTCGGGCT	394
Qy	317	TTTTCCGACACAGTTCGGGATGTTTCAGCCGGAAGCGCATCAGCAACCCGAAACATACCGG	376
Db	393	TTTTCCGACACAGTTCGGGATGTTTCAGCCGGAAGCGCATCAGCAACCCGAAACATACCGG	334
Qy	377	CGACAGCCGGAACCTGCGTGCAGATTTAATGACAGCGGTGCGGCGTGGGATA	436
Db	333	CGACAGCCGGAACCTGCGTGCAGATTTAATGACAGCGGTGCGGCGTGGGATA	274
Qy	437	TTAGTCAGGAGGACGGGTATCTGCTGGATGCCGAGAAATGGACATGGATACCCCG	496
Db	273	TTAGTCAGGAGGACGGGTATCTGCTGGATGCCGAGAAATGGACATGGATACCCCG	214
Qy	497	TGAGTTACCCGGGGCGGCTCGTTTCATTTCAGTTTTTGAACCCGTGGAGGACGGCA	556
Db	213	TGAGTTACCCGGGGCGGCTCGTTTCATTTCAGTTTTTGAACCCGTGGAGGACGGCA	154
Qy	557	GACTCGGGTGCAATGTGTTTTTACACGCTGATGGAGCAGATGAAGATGTCGACACGCT	616
Db	153	GACTCGGGTGCAATGTGTTTTTACACGCTGATGGAGCAGATGAAGATGTCGACACGCT	94
Qy	617	GCAGAACACCGAGCT	631
Db	93	GCAGAACACCGAGCT	79

RESULT 15	791 bp	mRNA	linear	EST 09-JUL-2003
CB520716/c	CB520716	NIH BMAP G10	Mus musculus	cdna clone
LOCUS	UI-M-G10-cej-j-06-0-UI.r1			
DEFINITION				

IMAGE:6840391 5', mRNA sequence.
ACCESSION CB520716
VERSION CB520716.1 GI:29354071
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .791
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840391"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpf"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 78.4%; Score 554.4; DB 6; Length 791;
Best Local Similarity 99.6%; Pred. No. 5e-155;
Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB 568 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 509
QY 105 TGGCGCTATCTGGGATCGGGAGGAAGAGCCGCTGCTTTTCCGGGAGGTGAAGCG 164
DB 508 TGGCGCTATCTGGGATCGGGAGGAAGAGCCGCTGCTTTTCCGGGAGGTGAAGCG 449
QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTCATTTGAGCGAAAGCGCACG 224
DB 448 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTCATTTGAGCGAAAGCGCACG 389
QY 225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTGTT 284
DB 388 TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTGTT 329
QY 285 CAGGCCACTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGCTCAGC 344

DB 328 CAGGCCACTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGCTCAGC 269
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGGAACTGCCGTGCCGGTGTG 404
DB 268 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGGAACTGCCGTGCCGGTGTG 209
QY 405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACCTCAGCGAGGACGGGTATCTCTGGC 464
DB 208 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACCTCAGCGAGGACGGGTATCTCTGGC 149
QY 465 TGGATCGCGCAGAAATCGACATCGGATACCCCGTGAGTTACCCGCGGGCGGCTCGTTC 524
DB 148 TGGATCGCGCAGAAATCGACATCGGATACCCCGTGAGTTACCCGCGGGCGGCTCGTTC 89
QY 525 ATTCAACGTTTTTGAACCCGTCGAGGACGGGACAGACTCGCGTGCAAAATGTGTTTAC-AG 583
DB 88 ATTCAACGTTTTTGAACCCGTCGAGGACGGGACAGACTCGCGTGCAAAATGTGTTTACNAG 29
QY 584 CGTATCGGACGAGATGAAGATGCTCGAC 611
DB 28 CGTATCGGACGAGATGAAGATGCTCGAC 1

Search completed: October 13, 2005, 06:19:24
Job time : 27488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 14:15:25 ; Search time 4777 Seconds
(without alignments)
242.170 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggatcccatgcgtcaatttt.....aaattgacgcatggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.2	80.1	46819	3	US-09-453-702B-72 Sequence 72, Appl
2	561.4	79.4	22306	3	US-09-453-702B-251 Sequence 251, Appl
3	223	31.5	34063	3	US-09-453-702B-96 Sequence 96, Appl
4	73.8	10.4	5194	3	US-08-844-274-16 Sequence 16, Appl
5	73.8	10.4	5194	4	US-09-598-421-16 Sequence 16, Appl
6	73.4	10.4	5194	3	US-08-844-274-17 Sequence 17, Appl
7	73.4	10.4	5194	4	US-09-598-421-17 Sequence 17, Appl
8	67.8	9.6	5679	3	US-08-844-274-10 Sequence 10, Appl
9	67.8	9.6	5679	4	US-09-598-421-10 Sequence 10, Appl
10	67.8	9.6	6723	3	US-08-844-274-13 Sequence 13, Appl
11	67.8	9.6	6723	4	US-08-844-274-14 Sequence 14, Appl
12	67.8	9.6	6723	4	US-09-598-421-13 Sequence 13, Appl
13	67.8	9.6	6723	4	US-09-598-421-14 Sequence 14, Appl
14	67.8	9.6	7560	3	US-08-844-274-20 Sequence 20, Appl
15	67.8	9.6	7560	4	US-09-598-421-20 Sequence 20, Appl
16	67.8	9.6	9423	4	US-09-377-066-6 Sequence 6, Appl
17	64.4	9.1	6448	3	US-08-844-274-15 Sequence 15, Appl
18	64.4	9.1	6448	4	US-09-598-421-15 Sequence 15, Appl
19	63.8	9.0	2476	3	US-08-844-274-11 Sequence 11, Appl
20	63.8	9.0	2476	4	US-09-598-421-11 Sequence 11, Appl
21	39.8	5.6	888	4	US-09-252-991A-15689 Sequence 15689, A
22	39.8	5.6	888	4	US-09-252-991A-15722 Sequence 15722, A
23	39.8	5.6	2001	4	US-09-252-991A-15589 Sequence 15589, A
24	39.8	5.6	2475	4	US-09-252-991A-15758 Sequence 15758, A
25	39.2	5.5	1734	6	Patent No. 5352575
26	39.2	5.5	1734	6	Patent No. 5352575-8
27	35.4	5.0	209210	4	US-09-949-016-15094 Sequence 15094, A

28	33.8	4.8	1254	4	US-09-252-991A-3154 Sequence 3154, Ap
29	33.8	4.8	1692	4	US-09-252-991A-2816 Sequence 2816, Ap
30	33.2	4.7	4403765	3	US-09-103-840A-2 Sequence 2, Appl
31	33.2	4.7	4411529	3	US-09-103-840A-1 Sequence 1, Appl
32	33	4.7	408	4	US-09-902-540-7848 Sequence 7848, Ap
33	33	4.7	449	4	US-09-702-705-1011 Sequence 1011, Ap
34	33	4.7	449	4	US-09-736-457-1011 Sequence 1011, Ap
35	33	4.7	449	4	US-09-614-124B-1011 Sequence 1011, Ap
36	33	4.7	449	4	US-09-671-325-1011 Sequence 1011, Ap
37	33	4.7	449	4	US-09-658-824-1011 Sequence 1011, Ap
38	33	4.7	5125	4	US-09-902-540-785 Sequence 785, App
39	33	4.7	24754	4	US-09-902-540-1230 Sequence 1230, Ap
40	32.8	4.6	1689	4	US-09-799-451-573 Sequence 573, App
41	32.8	4.6	2823	4	US-09-270-767-1868 Sequence 1868, Ap
42	32.8	4.6	2823	4	US-09-270-767-17150 Sequence 17150, A
43	32.8	4.6	3468	1	US-07-951-715A-2 Sequence 2, Appl
44	32.8	4.6	3468	2	US-08-459-448A-2 Sequence 2, Appl
45	32.8	4.6	3468	3	US-08-459-595A-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-453-702B-72
; Sequence 72, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46819
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-453-702B-72
Query Match 80.1%; Score 566.2; DB 3; Length 46819;
Best Local Similarity . 97.8%; Pred. No. 6.6e-176;

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI

COUNTRY: US
ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 34063

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 96:

US-09-453-702B-96

Query Match 31.5%; Score 223; DB 3; Length 34063;

Best Local Similarity 61.18; Pred. No. 2.6e-62;

Matches 361; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```
QY 51 CTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCCGCCAAGCTGGCGC 110
DB 14695 CTGCATAAGGATCACAATGTCGGGCATATGTTCTTATCAGTCACGTCGCGAATGGCGC 14636
QY 111 TATCTGGGATCGGGAGGAGAACGCCGTCCTTTTCCCGCGAGGTTGAAGCGGCATGG 170
DB 14635 TGGCTGGGGATGCGGGAGACCGCAGCAAAAGCTTGTCTGATGAGGTGGAGCGGCGCTGG 14576
QY 171 AAAGAGTTTCCGAGGATGACTGCTGCTGCTGATTGACGTTGAGCGAAGAACGACGTTTACC 230
DB 14575 TCGAATACGCCGAAGGATGTCGTGGGAGATGACGTTGGAAGAAACGACGTTCCAG 14516
QY 231 ATGATGATTCGGGAAGGTGTGGCCATGCGCAGCCCTTTTAAACGGTGAACCTGTTTCGTTACGGCC 290
DB 14515 GAATTTATCCGTGAAGGTGTGGGGGTTTCATCGGTTTAAACGGGAATCTTTGTGACGCG 14456
QY 291 ACCTGGGATACCAATGTCGTGCGGCTTTTCGGGACACAGTTCCGGATGGTCAGCCCGAAG 350
DB 14455 GTCTGGGATACGGAAACCAACGACGATTTTCGTACGGGTTTTTAAAGCCGTGAGTCCGAAA 14396
QY 351 CGCATCAGCAACCGAACAATACCGGCGACGAGCGGAACTCCGTCGCGGTGTGCAGATT 410
DB 14395 CGGGTGACACCGCCAGGACGCGTATGGGGAACCGTTTCTGCGGGCGGGGTGGAGGTC 14336
QY 411 AATCAGACGCGTGGCGCTGGGATATTAAGTCAGCGAGGACGGGTATCTCTGGCTGGATG 470
DB 14335 GATCGATATGCGCGTGGTGGTACCATATCTGTGAGGATGATTTTCTCTCGTCCGGG 14276
QY 471 CCGCAGAAATGACATGATACCCCGTGAAGTTACCCGCGCGGCGCGCTCGTTCAATCAC 530
DB 14275 AGTGGACGATGGGAACGATCCCGCGTGAACTTCCACCCGCGCGTCCGGCCATGCTGCAT 14216
```

```
QY 531 GTTTTGAACCCGTGGAGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACACGGTGATG 590
DB 14215 ATTTTCGAGCCGCTGGAGGACGGGCGAGACCCGTTGGGGCCCAACCAAGTTTACACGGTCATG 14156
QY 591 GAGCAGATGAAGATGCTCGACACGCTGCGAGAACGACGACGCTAGATTAAACC 641
DB 14155 GAACGCTGAAGATGCTCGATTCTCCCTGCAGGCAACACAGCTTCAGTCGGGCC 14105
```

RESULT 4

US-08-844-274-16/c

; Sequence 16, Application US/08844274B

; Patent No. 6218185

; GENERAL INFORMATION:

; APPLICANT: Fraser Jr., Malcom J.

; APPLICANT: Shirk, Paul D.

; APPLICANT: Elick, Teri A.

; APPLICANT: Perera, Omathage

; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System

; TITLE OF INVENTION: for Insects

; FILE REFERENCE: 0148.96

; CURRENT APPLICATION NUMBER: US/08/844,274B

; EARLIER FILING DATE: 1997-04-18

; EARLIER APPLICATION NUMBER: 60/016,234

; EARLIER FILING DATE: 1996-04-19

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0 - beta

; SEQ ID NO 16

; LENGTH: 5194

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: plasmid

; OTHER INFORMATION: IFF2B/Xpuc18.1

US-08-844-274-16

Query Match 10.4%; Score 73.8; DB 3; Length 5194;

Best Local Similarity 83.2%; Pred. No. 2.3e-13;

Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
QY 605 GCTCGACACGCTGCAGAACACGACGCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 664
DB 2955 GCTCGGTATCCCGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896
QY 665 GTACGTTAAAGATAAATCATGCGTAAATTCAGCATGGGAT 705
DB 2895 GTACGTTAAAGATAAATCATGCGTAAATTCAGCATGGT 2855
```

RESULT 5

US-09-598-421-16/c

; Sequence 16, Application US/09598421

; Patent No. 6551825

; GENERAL INFORMATION:

; APPLICANT: Fraser Jr., Malcom J.

; APPLICANT: Shirk, Paul D.

; APPLICANT: Elick, Teri A.

; APPLICANT: Perera, Omathage

; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System

; TITLE OF INVENTION: for Insects

; FILE REFERENCE: 0148.96

; CURRENT APPLICATION NUMBER: US/09/598,421

; CURRENT FILING DATE: 2000-06-19

; PRIOR APPLICATION NUMBER: 60/016,234

; PRIOR FILING DATE: 1996-04-19

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0 - beta

; SEQ ID NO 16

; LENGTH: 5194

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

```
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
US-09-598-421-16

Query Match          10.4%; Score 73.8; DB 4; Length 5194;
Best Local Similarity 83.2%; Pred. No. 2.3e-13;
Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 605 GCTCGACAGCGCTGACAGACACGCGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGCAC 664
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2955 GCTCGGTACCCGGGGATCGGATCCCTCTAGATTAAACCCCTAGAAAGATAATCATATTGTGCAC 2896

QY 665 GTACGTTAAAGATAATCATGCGTAAATAATTGACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2895 GTACGTTAAAGATAATCATGCGTAAATAATTGACGCATGTGTT 2855

RESULT 6
US-08-844-274-17/c
; Sequence 17, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
US-08-844-274-17

Query Match          10.4%; Score 73.4; DB 3; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 615 CTGCAGAACACGCGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAA 674
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4773 CTCTAGAGGGATCCTCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAA 4714

QY 675 GATAATCATGCGTAAATAATTGACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4713 GATAATCATGCGTAAATAATTGACGCATGTGTT 4683

RESULT 7
US-09-598-421-17/c
; Sequence 17, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-08-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19

; OTHER INFORMATION: Description of Artificial Sequence:plasmid
US-09-598-421-17

Query Match          10.4%; Score 73.4; DB 4; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 605 GCTCGACAGCGCTGACAGACACGCGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGCAC 664
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2955 GCTCGGTACCCGGGGATCGGATCCCTCTAGATTAAACCCCTAGAAAGATAATCATATTGTGCAC 2896

QY 665 GTACGTTAAAGATAATCATGCGTAAATAATTGACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2895 GTACGTTAAAGATAATCATGCGTAAATAATTGACGCATGTGTT 2855

RESULT 8
US-08-844-274-10
; Sequence 10, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p3E1.2 delta
US-08-844-274-10

Query Match          9.6%; Score 67.8; DB 3; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.3e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAATTTG 694
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1080 TTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1140 ACGCATGTGTT 1150

RESULT 9
US-09-598-421-10
; Sequence 10, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
```

```
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3el.2 delta
; OTHER INFORMATION: TRL
US-09-598-421-10

Query Match          9.6%; Score 67.8; DB 4; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACCGATGGGAT 705
Db 1140 ACCGATGTGTT 1150

RESULT 10
US-08-844-274-13
; Sequence 13, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; PRIOR FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-08-844-274-13

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACCGATGGGAT 705
Db 1140 ACCGATGTGTT 1150

RESULT 11
US-08-844-274-14/c
; Sequence 14, Application US/08844274B
```

```
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; PRIOR FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-08-844-274-14

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

QY 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 12
US-09-598-421-13
; Sequence 13, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-09-598-421-13

Query Match          9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150
```

```
RESULT 13
US-09-598-421-14/c
; Sequence 14, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omathchage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-09-598-421-14

Query Match          9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

Qy 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 14
US-08-844-274-20/c
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omathchage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-08-844-274-20

Query Match          9.6%; Score 67.8; DB 3; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694

Qy 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 15
US-09-598-421-20/c
; Sequence 20, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omathchage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-09-598-421-20

Query Match          9.6%; Score 67.8; DB 4; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694

Qy 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

Search completed: October 13, 2005, 07:39:19
Job time : 4784 secs
```